International Baccalaureate®
Baccalauréat International
Bachillerato Internacional

Extended essay cover

Candidates must complete this page and then give this cover and their final version of the extended essay to their supervisor.

Candidate session number			
Candidate name			
School name			
Examination session (May or November)	MAY	Year	2015

Diploma Programme subject in which this extended essay is registered: MATHEMATICS

(For an extended essay in the area of languages, state the language and whether it is group 1 or group 2.)

Title of the extended essay: GENERATING ARBITRARY UNIFORM PROBABILITY DISTRIBUTIONS

Candidate's declaration

This declaration must be signed by the candidate; otherwise a mark of zero will be issued.

The extended essay I am submitting is my own work (apart from guidance allowed by the International Baccalaureate).

I have acknowledged each use of the words, graphics or ideas of another person, whether written, oral or visual.

I am aware that the word limit for all extended essays is 4000 words and that examiners are not required to read beyond this limit.

This is the final version of my extended essay.

Candidate's signature: Date: 29/01/2015

Supervisor's report and declaration

The supervisor must complete this report, sign the declaration and then give the final version of the extended essay, with this cover attached, to the Diploma Programme coordinator.

Name of supervisor (CAPITAL letters) [redacted]

Please comment, as appropriate, on the candidate's performance, the context in which the candidate undertook the research for the extended essay, any difficulties encountered and how these were overcome (see page 13 of the extended essay guide). The concluding interview (viva voce) may provide useful information. These comments can help the examiner award a level for criterion K (holistic judgment). Do not comment on any adverse personal circumstances that may have affected the candidate. If the amount of time spent with the candidate was zero, you must explain this, in particular how it was then possible to authenticate the essay as the candidate's own work. You may attach an additional sheet if there is insufficient space here.

But I don't! The explanations should appear [in] the es[say]

[redacted] is an exceptional Mathematician. One of the best I have ever encountered. I was far more daunted by his EE than he was and I set him the task of ensuring that I understood everything by the end. He did this!

He lacks[…] but the explanation is in the essay

The Mathematics is at a very high level but he clearly understands it all and can explain any aspect of the EE. He generated the problem entirely independently and had to be persuaded to focus his essay as he was juggling many strands initially. – He needed more help here

He didn't start be explaining ba[…]

He was also encouraged to ensure the stages in his work were clear as what is obvious to him is not always obvious to others!!

This declaration must be signed by the supervisor; otherwise a mark of zero will be issued.

I have read the final version of the extended essay that will be submitted to the examiner.

To the best of my knowledge, the extended essay is the authentic work of the candidate.

As per the section entitled "Responsibilities of the Supervisor" in the EE guide, the recommended number of hours spent with candidates is between 3 and 5 hours. Schools will be contacted when the number of hours is left blank, or where O hours are stated and there lacks an explanation. Schools will also be contacted in the event that number of hours spent is significantly excessive compared to the recommendation.

The supervisor has failed to ensure the candidates shows the understanding mentioned above in the essay.

I spent $\boxed{4}$ hours with the candidate discussing the progress of the extended essay.

Supervisor's signature: [redacted] Date: 27/01/2015

Assessment form (for examiner use only)

Candidate session number		

Criteria	Achievement level				
	Examiner 1	maximum	Examiner 2	maximum	Examiner 3
A research question	1	2		2	
B introduction	1	2		2	
C investigation	3	4		4	
D knowledge and understanding	2	4		4	
E reasoned argument	3	4		4	
F analysis and evaluation	2	4		4	
G use of subject language	3	4		4	
H conclusion	2	2		2	
I formal presentation	3	4		4	
J abstract	2	2		2	
K holistic judgment	3	4		4	
Total out of 36	25				

Name of examiner 1: ______________________ Examiner number: __________
(CAPITAL letters)

Name of examiner 2: ______________________ Examiner number: __________
(CAPITAL letters)

Name of examiner 3: ______________________ Examiner number: __________
(CAPITAL letters)

IB Assessment Centre use only: B: ✓

IB Assessment Centre use only: A: __________

International Baccalaureate

Extended Essay

Generating Arbitrary Uniform Probability Distributions

Word count: 3258

Abstract

This paper presents a general method to generate points distributed uniformly across any arbitrary spatial region or boundary. Although computers can generate one-dimensional uniform distributions quite effectively, uniform distributions across more complex regions in higher dimensions cannot directly be generated. So, this paper aims to explore the process of producing points in an arbitrary uniform distribution beginning with the generation of only one-dimensional random variables. RQ

An initial observation is made through exploring the use of parametric equations over uniformly distributed parameters to directly produce distributions, which fails to a lack of uniformity over the resultant region. A combination of multivariate probability theory, vector calculus, and mathematical reasoning leads to a solution applicable to continuous and differentiable regions, which is then extended to boundary cases. Upon examination of a specific complex case with nondifferentiable points, a geometric approach is then used to improve on and generalize the result. method

The main result of this study is summarized as a four-step procedure that can be applied to generate uniform distributions over generalized regions and boundaries. As an exploration, this paper also evaluates possible problems with the method while providing some useful general solutions that can be used in a variety of distribution problems. The result of this investigation offers the possibility of generalizing other one-dimensional distributions to higher dimensions in the future. Conc.

(223 words)

Clear.

Contents

1 Introduction

Uniform probability distributions prove to be a highly useful aspect of mathematics, with applications in topics ranging from risk assessment to modeling. Notably, the uniform distribution is a commonly used prior distribution (representing information known of an uncertain parameter before any information is gained) in Bayesian inference (Gelman 6-7). Additionally, the uniform distribution often forms a precursor to generating a more complex distribution, as the most commonly random number generation algorithms used today are optimized to generate values with equal probabilities. These algorithms, known as pseudorandom number generators (PRNGs), include the Mersenne Twister, Lehmer PRNG, and linear congruential generators, amongst many others.

What?

|| example?

One can easily extend these one-dimensional uniform distributions into higher dimensions by taking multiple random values as additional Cartesian coordinates, generating distributions over squares, cubes, tesseracts, and hypercubes. However, uniform distributions over arbitrarily (non-hypercube) shaped regions, such as the position of any water molecule in a jug, also cannot be generated directly with PRNGs. **For example,** Monte Carlo modeling of photon propagation notably requires the generation of uniform points across the surface of a sphere, useful for statistical physics and engineering models (Penzov et al.). Interestingly, these arbitrary uniform distributions are already quite difficult to generate even for shapes as intuitively ~~simple as a sphere;~~ this problem provided the motivation behind an exploration of the topic. This essay will outline a general method to generate these arbitrary uniform distributions, constructing a solution through multivariate probability theory, vector calculus, and geometry while exploring some potential mathematical pitfalls of the process.

|| Show!

What exactly is a uniform distribution. Start here! EXPLAIN

Discrete or continuous?

2 Naive generation

2.1 Sphere distributions

These are essential to the essay but not clearly defined.

Parameterization offers a powerful way to express abstract mathematical objects in terms of known variables. For example, a unit sphere can be generated with its well-known parametric equations in spherical coordinates $x = \rho\cos\theta\sin\phi$, $y = \rho\sin\theta\sin\phi$, and $z = \rho\cos\phi$, where $0 \le \rho < 1$, $0 \le \phi < \pi$, and $0 \le \theta < 2\pi$ The immediate, intuitive approach to generate a uniform distribution across this sphere is to randomize the parameters ρ, θ, and ϕ uniformly for each point. The result of 10^5 of these points is illustrated in Figure 1, and the same points viewed through the three axis planes are shown in Appendix A, Table 2. 10^5 was chosen as a sufficiently ~~large number of random~~ points to fully visualize the distribution that is generated using uniform parameter distributions.

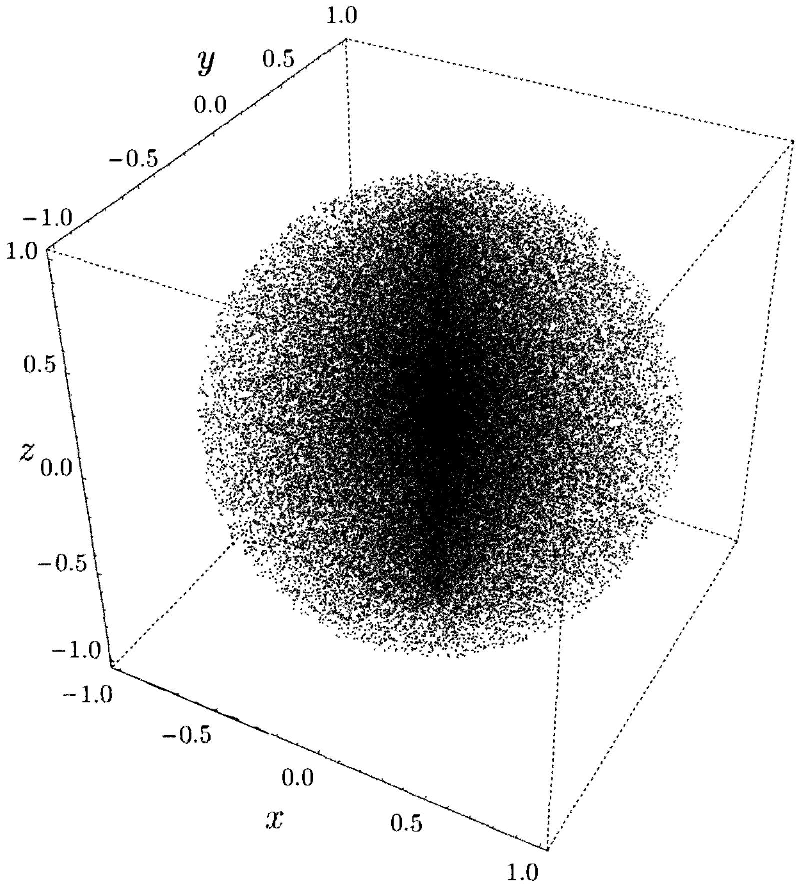
What are? How was this done?

Justify these.

Should be here

Inappropriate use of appendices. (I)

Figure 1: Sphere generated by uniform parameters — How?



Clearly, this method of generation yields a distribution that does not visually appear uniform. The points qualitatively appear more densely distributed about the vertical axis of the sphere along the z-axis, with the greatest density of points around the sphere's center. Intuitively, this result should be quite predictable, especially when considering the consequences of generating ρ uniformly. While we generated a theoretically equal number of points between $0 \leq \rho < 0.5$ and $0.5 \leq \rho < 1$, the first range encompasses a volume of $4\pi(0.5)^3/3 \approx 0.524$, while the second range encompasses a volume of $4\pi(1^3 - 0.5^3)/3 \approx 3.665$. Evidently, an equal number of points in both of these regions will result in a denser distribution across the first.

yes but not explained at all how the points were generated.

2.2 Jacobian determinants

Still not defined.

To describe this distortion of the uniform distribution mathematically, the Jacobian determinant is used; it is the determinant of the Jacobian matrix of an n-dimensional transformation $\mathbf{x} : \mathbb{R}^n \to \mathbb{R}^n$ of functions $x_1, x_2, ..., x_n$ over parameters $\mathbf{t} = (t_1, t_2, ..., t_n)$, and quantifies the expansion or contraction in volume around every point. The Jacobian matrix is defined as:

$$\frac{\partial \mathbf{x}}{\partial \mathbf{t}} = \begin{bmatrix} \frac{\partial x_1}{\partial t_1} & \cdots & \frac{\partial x_1}{\partial t_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial t_1} & \cdots & \frac{\partial x_n}{\partial t_n} \end{bmatrix}$$

In the case of the sphere above, its Jacobian determinant is:

Not clear at all why this is so — or why it is needed.

$$\begin{aligned} \det \frac{\partial(x,y,z)}{\partial(\rho,\phi,\theta)} &= \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \phi} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \phi} & \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial \rho} & \frac{\partial z}{\partial \phi} & \frac{\partial z}{\partial \theta} \end{vmatrix} \\ &= \begin{vmatrix} \cos\theta\sin\phi & \rho\cos\theta\cos\phi & -\rho\sin\theta\sin\phi \\ \sin\theta\sin\phi & \rho\sin\theta\cos\phi & \rho\cos\theta\sin\phi \\ \cos\phi & -\rho\sin\phi & 0 \end{vmatrix} \\ &= \rho^2\cos^2\theta\sin\phi\cos^2\phi + \rho^2\sin^2\theta\sin^3\phi \\ &\quad + \rho^2\cos^2\theta\sin^3\phi + \rho^2\sin^2\theta\sin\phi\cos^2\phi \\ &= \rho^2\sin\phi\,(\sin^2\theta + \cos^2\theta)(\sin^2\phi + \cos^2\phi) \\ &= \rho^2\sin\phi \end{aligned}$$

?

?

Why?

When the Jacobian determinant is a positive value greater than 1, it indicates that the volume around a point is expanded upon the transformation; as a result, uniformly generated points are spaced further apart after transformation. The opposite is true for positive values less than 1. For negative values, the Jacobian ~~determinant~~ describes similar behavior, but the orientation of these volumes is reversed (Amidror 310). Don't quote. Explain

The Jacobian determinant of our sphere increases as ρ increases, moving towards the sphere boundary, or when ϕ is around $\pi/2$, around the sphere's equator. Hence, the expansion in volume is greatest at these areas, and the generated points are spaced furthest apart. This demonstrates that the 'naive' uniform parameter approach does not generate a uniform distribution over the sphere. ✓ Yes, but why the Jacobian determinant demonstrates this is NOT made clear

3 Uniform generation

3.1 Extension of distributions to higher dimensions

Although the uniform parameter approach does not directly yield the desired uniform sphere distribution, directly adjusting how the parameters are generated should result in a uniform

You still haven't defined your single variable uniform distr

distribution. This requires an extension of probability density functions for a single random variable (RV) to multiple. The definitions below are adaptations of those of Billingsley (261-262).

Definition 3.1. *For continuous RVs* $\mathbf{X} = (X_1, X_2, ..., X_n)$*, the joint probability density function* $f_{\mathbf{X}} \geq 0$ *represents the relative probability that* $\mathbf{X}$ *lies in a set* A *over values* $x_1, x_2, ..., x_n$*, and satisfies:*

$$\begin{aligned} P(\mathbf{X} \in A) &= \int_A f_{\mathbf{X}}(x_1, x_2, ..., x_n)\, dx_1\, dx_2 \ldots dx_n \\ &= 1 \qquad \textit{if } A \textit{ includes all possible } \mathbf{X}. \end{aligned}$$

Not clear what this means.

Definition 3.2. *The marginal probability density function* f_{X_i} *is equivalent to the joint probability density considering only* X_i *alone, and is obtained through integrating the joint probability density function with respect to all variables except* x_i*, satisfying:*

$$f_{X_i}(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, x_2, ..., x_n)\, dx_1 \ldots dx_{i-1}\, dx_{i+1} \ldots dx_n$$

Why? Justify.

Definition 3.3. *The marginal cumulative distribution function* F_{X_i} *satisfies:*

$$F_{X_i}(x_i) = \int_{-\infty}^{x_i} f_{X_i}(x_i)\, dx_i$$

OK.

Not explained why the Jacobian determinant describes this

3.2 Applying distribution adjustments

While uniformly distributed parameters do not generate uniform distributions in the general case due to the distortion of volumes described by the Jacobian determinant, one might expect that well-chosen non-uniformly distributed parameters can be transformed into the uniform distribution we desire. So, the adjustments required should be functions applied to the generated uniform RVs $\psi_i \in [0, 1)$ which produce that specific non-uniform parameter distribution, counteracting the distortion caused by the parametric equations. ?

what are these?

If we treat the Jacobian determinant as a joint probability density function (which must be first normalized) for random variables R, Φ, Θ, then the functions r, ϕ, θ to apply which correspond to those random variables should be equal to the inverses of the marginal cumulative distribution functions for each parameter. The proof of the validity of this method will follow in Section 4. Note, however, that the Jacobian determinant is in itself not a probability density function, as it often fails to satisfy the requirement of Definition 3.1 that the integral of the density over all possible values is equal to 1. So, the Jacobian determinant needs to be multiplied by some normalizing constant that produces a distribution function with this property. First:

You haven't explained what such an integral means.

$$\int_0^{2\pi}\int_0^{\pi}\int_0^{1}\left|\frac{\partial(x,y,z)}{\partial(\rho,\phi,\theta)}\right| d\rho\, d\phi\, d\theta = \int_0^{2\pi}\int_0^{\pi}\int_0^{1}\rho^2\sin\phi\, d\rho\, d\phi\, d\theta$$
$$= \int_0^{2\pi}\int_0^{\pi}\left[\frac{1}{3}\rho^3\sin\phi\right]_0^1 d\phi\, d\theta$$
$$= \int_0^{2\pi}\int_0^{\pi}\frac{1}{3}\sin\phi\, d\phi\, d\theta$$
$$= \int_0^{2\pi}\left[-\frac{1}{3}\cos\phi\right]_0^{\pi} d\theta$$
$$= \int_0^{2\pi}\frac{2}{3}\, d\theta$$
$$= \frac{4\pi}{3}$$

So, $3/4\pi$ needs to be ~~multiplied to~~ the Jacobian determinant to normalize it into a valid density function. Hence, $\dfrac{3}{4\pi}\rho^2\sin\phi$ is the joint probability density function. Then, each of the marginal cumulative ~~distributions~~ distributions can be directly calculated:

A pdf needs to be more clearly defined than this.

$$F_R(\rho) = \int_0^{\rho}\int_0^{2\pi}\int_0^{\pi}\frac{3}{4\pi}\rho^2\sin\phi\, d\phi\, d\theta\, d\rho$$
$$= \int_0^{\rho}3\rho^2\, d\rho$$
$$= \rho^3$$
$$r(\psi_1) = \sqrt[3]{\psi_1}$$

What does this mean?

$$F_\Phi(\phi) = \int_0^{\phi}\int_0^{2\pi}\int_0^{1}\frac{3}{4\pi}\rho^2\sin\phi\, d\rho\, d\theta\, d\phi$$
$$= \int_0^{\phi}\frac{1}{2}\sin\phi\, d\phi$$
$$= \frac{1}{2}(1-\cos\phi)$$
$$= \sin^2\frac{\phi}{2}$$
$$\phi(\psi_2) = 2\arcsin\sqrt{\psi_2}$$

and this?

$$F_\Theta(\theta) = \int_0^{\theta}\int_0^{\pi}\int_0^{1}\frac{3}{4\pi}\rho^2\sin\phi\, d\rho\, d\phi\, d\theta$$

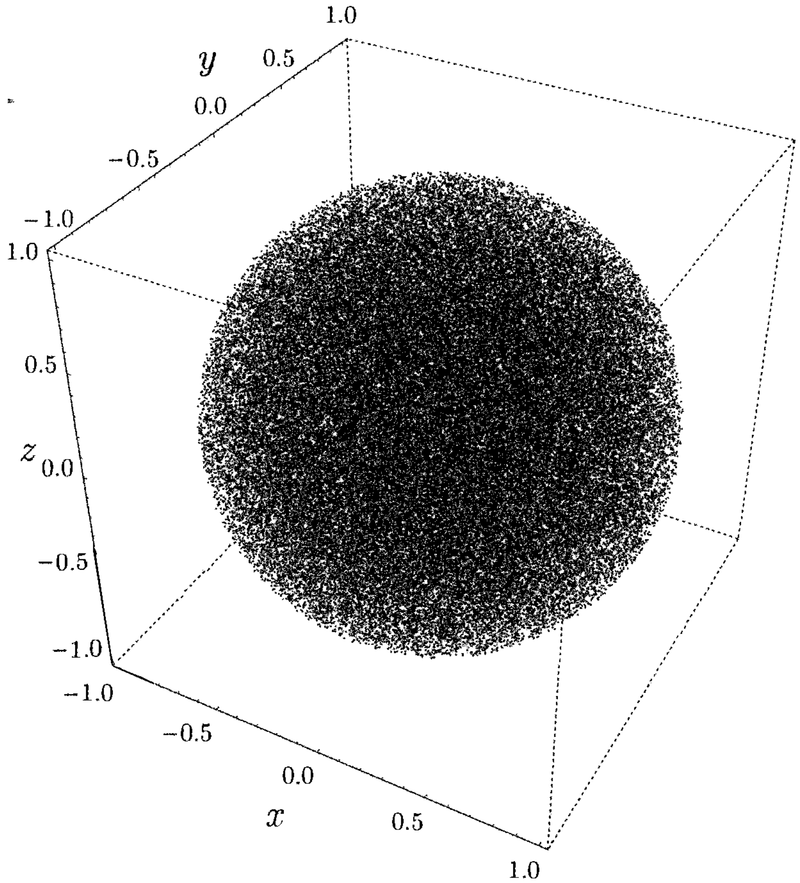
$$
\begin{aligned}
&= \int_0^{\theta} \frac{1}{2\pi}\, d\theta \\
&= \frac{\theta}{2\pi} \\
\theta(\psi_3) &= 2\pi\psi_3
\end{aligned}
$$

And this ?

Note that these transformations are nearly identical those used in the Monte Carlo cosine weighted random sampling method for uniform point sets on a sphere surface used in image rendering (Penzov et al.). The result of 10^5 points generated by these transformations is illustrated in Figure 2, and the same points projected onto the three axis planes is also shown in contrast with the naive distributions in Appendix A, Table 2. It is evident that the distribution visually appears quite uniform from all directions, largely contrasting with the results in Figure 1.

How ? No clear idea of how you did this. Or are these diagrams taken from one of your sources? You don't show how you do it.

Figure 2: Sphere generated by adjusted parameters



4 Generalization of solutions

A more rigorous explanation of the validity of the solution to Section 3 is evidently necessary. While the general idea of applying the inverse cumulative distribution function is a common method of generating a desired distribution from a uniform in one dimension, it is not extremely clear how it can be extended into more dimensions. The proof presented can be outlined as a solution for the unknown transformations $t_1, t_2, ..., t_n$ involved in the progression of random variables illustrated in Figure 3.

Figure 3: Progression of variable transformations in generalized solution

Computer-generated random variables		Adjusted-distribution random variables		Region-distributed random variables
RVs Ψ_i with values ψ_i	$\xrightarrow{t_i}$	RVs T_i with values t_i	$\xrightarrow{x_i}$	RVs X_i with values x_i
Uniform across $[0, 1)$		Non-uniform across $[0, 1)$		Uniform across shape

? Not clear what this means. explain

4.1 Solution for spatial regions

Suppose our n-dimensional region S is defined through a system of n parametric equations with n parameters $\mathbf{x}(\mathbf{t}) = (x_1(\mathbf{t}), x_2(\mathbf{t}), ..., x_n(\mathbf{t}))$, with $\mathbf{t} = (t_1, t_2, ..., t_n) \in [0, 1)^n$. When producing a random point on S, we must generate n uniformly distributed random variates $\Psi = (\Psi_1, \Psi_2, ..., \Psi_n)$ across $[0, 1)^n$. We must find the transformations t_i for each of these random variates such that the joint density function $f_{\mathbf{X}}$ of independent random variates $\mathbf{X} = \mathbf{x}(\mathbf{t}(\mathbf{\Psi}))$ across S is constant. Thus, $f_{\mathbf{X}}(\mathbf{x}) = \alpha^{-1}$ for some nonzero constant α when $\mathbf{x} \in S$.

? Not clear.

Theorem 4.1. *The inverse function of t_i is equal to the marginal cumulative distribution function for T_i.*

Proof. The marginal probability density function f_{Ψ_i} for random variate Ψ_i satisfies:

$$P(a \leq \Psi_i < b) = \int_a^b f_{\Psi_i}(\psi_i)\, d\psi_i$$

Thus, upon applying the transformation to another random variate $T_i = t_i(\Psi_i)$:

$$P(t_i(a) \leq T_i < t_i(b)) = \int_{t_i(a)}^{t_i(b)} f_{\Psi_i}(\psi_i(t_i)) \frac{d\psi_i}{dt_i}\, dt_i$$

?

Therefore:

$$f_{T_i}(t_i) = f_{\Psi_i}(\psi_i(t_i)) \frac{d\psi_i}{dt_i}$$

It is known that Ψ_i is uniform across $[0, 1)$ and therefore $f_{\Psi_i}(\psi_i) = 1$. Then, as $t_i \in [0, 1)$:

$$\begin{aligned} f_{T_i}(t_i) &= \frac{d\psi_i}{dt_i} \\ F_{T_i}(t_i) &= \int_0^{\psi_i(t_i)} \frac{d\psi_i}{dt_i}\, dt_i \\ &= \psi_i(t_i) \\ t_i(\psi_i) &= F_{T_i}^{-1}(\psi_i) \end{aligned}$$

□

Using this theorem, we can extrapolate what is needed to obtain t_i. Suppose that R is a region in $[0, 1)^n$ which maps by the parametric equations $\mathbf{x} : R \to D$, where D is a subregion of S. By Definition 3.1, $f_{\mathbf{X}}$ satisfies by a variable transformation:

$$P(\mathbf{X} \in D) = \int_D f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} = \int_R \alpha^{-1} \left| \frac{\partial \mathbf{x}}{\partial \mathbf{t}} \right| d\mathbf{t}$$

The marginal density functions $f_{T_i}(t_i)$ thus satisfy by Definition 3.2:

$$f_{T_i}(t_i) = \alpha^{-1} \int_0^1 \cdots \int_0^1 \left| \frac{\partial \mathbf{x}}{\partial \mathbf{t}} \right| dt_1 \ldots dt_{i-1}\, dt_{i+1} \ldots dt_n$$

The marginal cumulative distribution follows from Definition 3.3:

$$F_{T_i}(t_i) = \alpha^{-1} \int_0^{t_i} \int_0^1 \cdots \int_0^1 \left| \frac{\partial \mathbf{x}}{\partial \mathbf{t}} \right| dt_1 \ldots dt_{i-1}\, dt_{i+1} \ldots dt_n\, dt_i \tag{1}$$

Hence we have obtained an expression for t_i^{-1}. For region problems in general, the approach to creating a uniform distribution will involve computing each F_{T_i} and inverting them to obtain t_i. With the property that $F_{T_i}(1) = 1$, the value of α can be determined:

$$\alpha = \int_0^1 \cdots \int_0^1 \left| \frac{\partial \mathbf{x}}{\partial \mathbf{t}} \right| dt_1 \ldots dt_n = \int_S d\mathbf{x} \tag{2}$$

Not clear.

never explained.

quite.

These rather daunting equations are only deceptively complex. In three dimensions, $f_{T_i}(t_i)$ is only a double integral and the Jacobian determinant may sometimes be independent with respect to one of the parameters. This has already been demonstrated in Section 3, where the Jacobian of a ~~sphere is independent~~ of the third parameter θ. In addition, the normalizing constant α is equivalent to the total volume of the region, which is relatively simple to calculate for common regions like the Platonic solids or ellipsoids.

Why? explain!

4.2 Solution for spatial boundaries

While Section 4.1 illustrates a generalized solution for regions, such as the region enclosed by a sphere, it does not apply to generating a boundary like the surface of the sphere. To solve the problem of uniform boundaries, note that boundaries in n dimensions can be expressed in terms of a corresponding region; suppose a parameterization $\mathbf{x}(\mathbf{t})$ of the space S exists such that setting the last parameter $t_n = 1$ obtains a boundary B of S. Then it follows from Equation 1 that instead of allowing $0 \leq t_n < 1$, it can instead be set to $t_n = 1$ to obtain a uniform distribution over B. Thus, the marginal cumulative distributions for a boundary can be instead described as (with $i \neq n$):

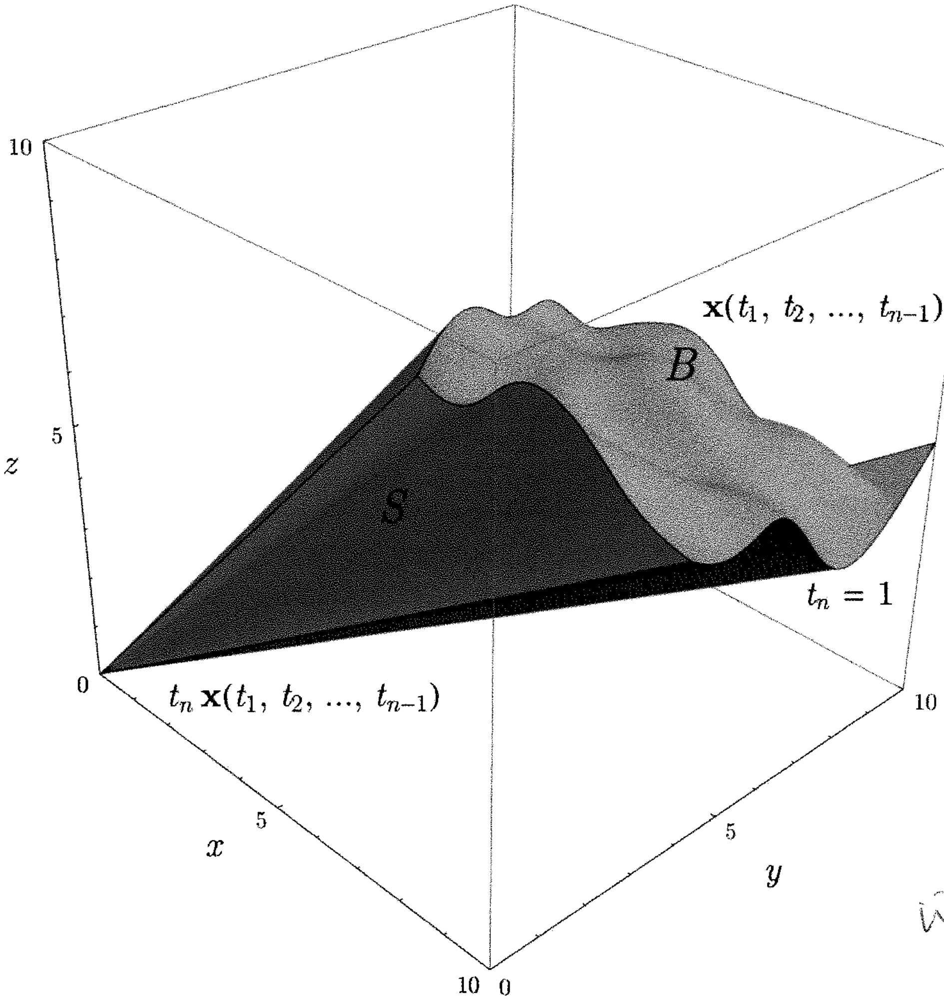
$$F_{T_i}(t_i) = \alpha^{-1} \int_0^{t_i} \int_0^1 \cdots \int_0^1 \left| \frac{\partial \mathbf{x}}{\partial \mathbf{t}} \right|_{t_n=1} dt_1 \ldots dt_{i-1}\, dt_{i+1} \ldots dt_{n-1}\, dt_i \tag{3}$$

For many of these boundaries we may expect initially that it is difficult to find an appropriate region which satisfies this property. Indeed, in the case that the boundary B over n dimensions we wish to generate is open, S may not be a finite region. For these situations, a parameterization should be obtained directly for B with $n-1$ parameters. For most boundaries, a multiplication by a new nth parameter is sufficient, generating an artificial region for which setting $t_n = 1$ obtains B, as illustrated in Figure 4. Some special cases may have better region parameterizations, like the paraboloid distribution; see Appendix B.2.2.

Again, inappropriate use of appendix (I).

Should be here.

Figure 4: Producing a region from a boundary



What does this show?

5 Constructing a solution algorithm

5.1 Coffee mug distributions

The solutions presented above may appear to cover most conceivable regions and boundaries. However, several assumptions have been made throughout the process. The final result for t_i^{-1} clearly shows the condition that the parameterization $\mathbf{x}(\mathbf{t})$ should exist and each of its components are differentiable with respect to all t_i. This is evidently a problem in polytopes, for which there may not be a satisfactory $\mathbf{x}(\mathbf{t})$ parameterization because of sharp edges and vertices. Indeed, many regions that are defined piecewise are nondifferentiable at specific points, and require an improved method to model with a uniform distribution; this provides a motivation to create an even more generalized algorithm for solving these cases.

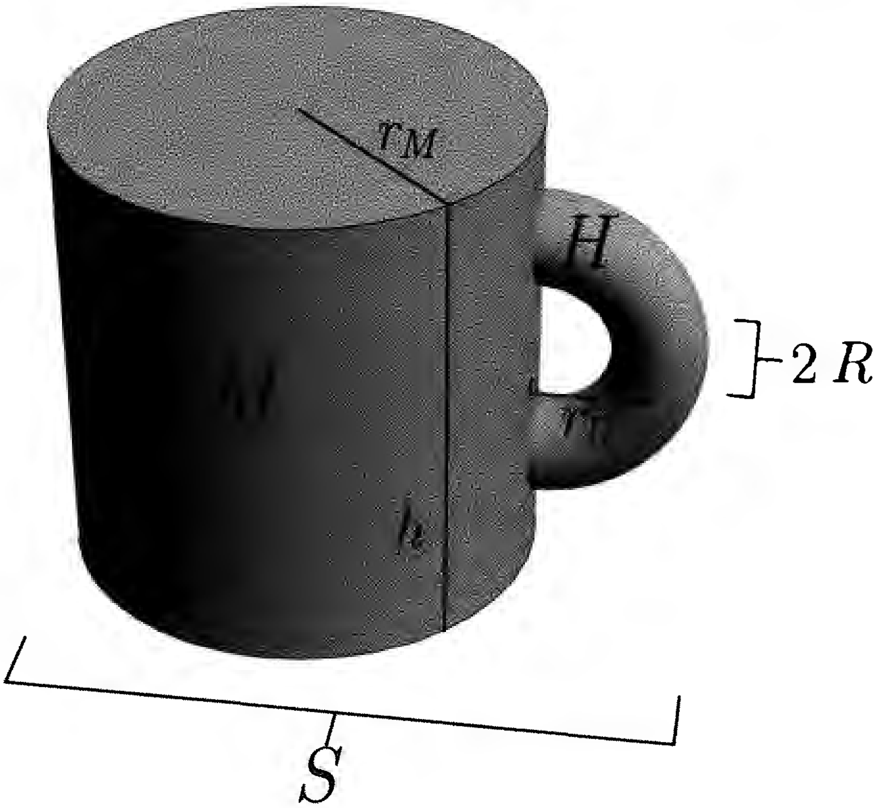
what?

example?

Many regions or boundaries for which a uniform distribution must be generated have parameterizations that are not differentiable everywhere. However, the problem may be simplified by considering that many nondifferentiable regions are essentially unions of multiple subregions, each of which is differentiable. Consider, for example, a coffee mug region S, illustrated in Figure 5; while the mug M and the handle H regions together approximate S and are nondifferentiable where they join, uniform distributions across the mug and the handle can be generated individually.

Figure 5: A coffee mug region

explain what this means. Why?



Suppose we want to generate a uniform distribution inside the coffee cup. Region M is a cylinder with parametric equations:

$$x_M(r,\theta,z) = r\cos\theta \qquad 0 \le r < r_M$$
$$y_M(r,\theta,z) = r\sin\theta \qquad 0 \le \theta < 2\pi$$
$$z_M(r,\theta,z) = z \qquad 0 \le z < h$$

Just quoted! No justification.

You ought to justify that these are valid parameters for M. You don't.

A shame you never clearly explained what these are

The Jacobian determinant is:

$$\left|\frac{\partial(x_M, y_M, z_M)}{\partial(r, \theta, z)}\right| = \begin{vmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix}$$
$$= r\cos^2\theta + r\sin^2\theta$$
$$= r$$

This is independent of both θ and z, so only r must be considered. Clearly, the volume $\alpha = \pi r_M^2 h$:

Not explained/justified why this gives a volume

$$F_R(r) = \frac{1}{\pi r_M^2 h}\int_0^r \int_0^{2\pi} \int_0^h r\, dz\, d\theta\, dr$$
$$= \frac{2}{r_M^2}\int_0^r r\, dr$$
$$= \frac{r^2}{r_M^2}$$
$$r(\psi_r) = r_M\sqrt{\psi_r}$$

?? Why?

The handle H can be parametrized as a half torus. Its center is located at $(r_M, 0, h/2)$, and it has the limitation $0 < R < r_H$. The half torus has the parameterization:

$$\begin{aligned} x_H(r,\theta,\phi) &= r_M + (r_H + r\cos\phi)\cos\theta & 0 &\le r < R \\ y_H(r,\theta,\phi) &= r\sin\phi & -\pi/2 &\le \theta < \pi/2 \\ z_H(r,\theta,\phi) &= h/2 + (r_H + r\cos\phi)\sin\theta & 0 &\le \phi < 2\pi \end{aligned}$$

Why? Justify

The Jacobian determinant is:

$$\left|\frac{\partial(x_H, y_H, z_H)}{\partial(r, \theta, \phi)}\right| = \begin{vmatrix} \cos\phi\cos\theta & -(r_H + r\cos\phi)\sin\theta & -r\sin\phi\cos\theta \\ \sin\phi & 0 & r\cos\phi \\ \cos\phi\sin\theta & (r_H + r\cos\phi)\cos\theta & -r\sin\phi\sin\theta \end{vmatrix}$$
$$= -r\sin\phi\,(r_H + r\cos\phi)(\sin\phi\sin^2\theta + \sin\phi\cos^2\theta)$$
$$- r\cos\phi\,(r_H + r\cos\phi)(\cos\phi\cos^2\theta + \cos\phi\sin^2\theta)$$
$$= -r\sin^2\phi\,(r_H + r\cos\phi) - r\cos^2\phi\,(r_H + r\cos\phi)$$
$$= -r\,(r_H + r\cos\phi)$$

The volume of this half torus is:

$$\begin{aligned}
\alpha &= \int_0^{2\pi}\int_{-\pi/2}^{\pi/2}\int_0^R -r\,(r_H + r\cos\phi)\,dr\,d\theta\,d\phi \\
&= -\int_0^{2\pi}\int_{-\pi/2}^{\pi/2}(R^2 r_H/2 + R^3\cos\phi/3)\,d\theta\,d\phi \\
&= -\pi\int_0^{2\pi}(R^2 r_H/2 + R^3\cos\phi/3)\,d\phi \\
&= -\pi^2 R^2 r_H
\end{aligned}$$

The Jacobian determinant is independent of θ.

$$\begin{aligned}
F_R(r) &= \frac{1}{\pi^2 R^2 r_H}\int_0^r\int_0^{2\pi}\int_{-\pi/2}^{\pi/2} r\,(r_H + r\cos\phi)\,d\theta\,d\phi\,dr \\
&= \frac{2}{R^2 r_H}\int_0^r r_H r\,dr \\
&= \frac{r^2}{R^2} \\
r(\psi_r) &= R\sqrt{\psi_r}
\end{aligned}$$

? How does this follow.

$$\begin{aligned}
F_\Phi(\phi) &= \frac{1}{\pi^{\cancel{2}}\cancel{R^2} r_H}\int_0^\phi\int_{-\pi/2}^{\pi/2}\int_0^R r\,(r_H + r\cos\phi)\,dr\,d\theta\,d\phi \\
&= \frac{1}{6\pi r_H}\int_0^\phi(3r_H + 2R\cos\phi)\,d\phi \\
&= \frac{3r_H\phi + 2R\sin\phi}{6\pi r_H}
\end{aligned}$$

Notably, this expression cannot simply be inverted into commonly used functions due to the separate trigonometric $2R\sin\phi$ and polynomial $3r_H\phi$ terms. This also suggests that it is necessary to check the validity of this cumulative distribution; for it to be valid, it must be monotonically increasing between $0 \le \phi < 2\pi$. Thus, we require for all ϕ:

$$\begin{aligned}
\frac{d}{d\phi}\frac{3r_H\phi + 2R\sin\phi}{6\pi r_H} &\geq 0 \\
3r_H + 2R\cos\phi &\geq 0 \\
3r_H - 2R &\geq 0 \qquad (r_H > R > 0)
\end{aligned}$$

why does this follow ?

This is trivially satisfied with $R < r_H$. Although we may expect a need to find an explicit expression for the inverse of F_Φ, it is not required in practice as it can be numerically calcu-

 ?

how? By whom?

lated. 10^5 points in each of the two distributions obtained are illustrated Figures 6 and 7, with $r_M = 1$, $h = 2$, $r_H = 0.5$, and $R = 0.2$, and projections of these distributions onto the three axis planes are illustrated in Appendix A, Figures 9 and 10. Should be here.

Figure 6: M distribution

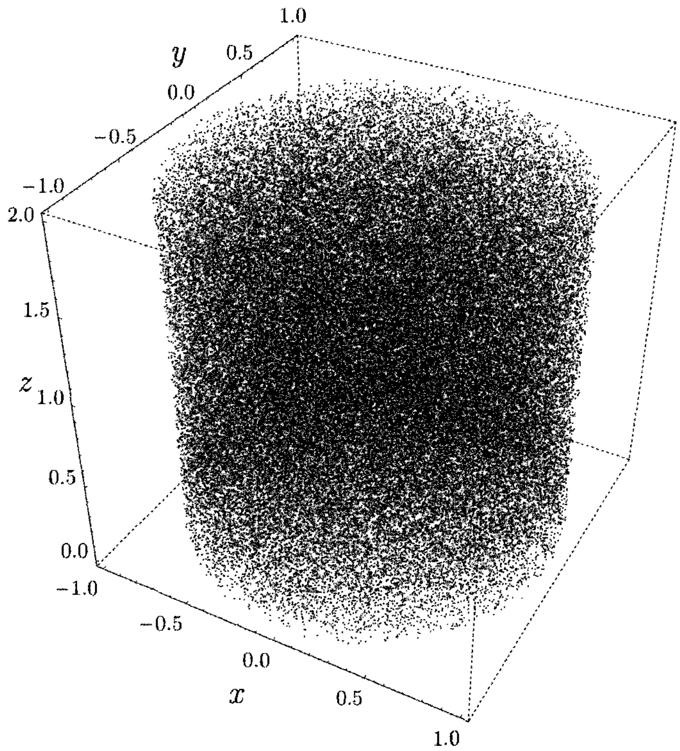
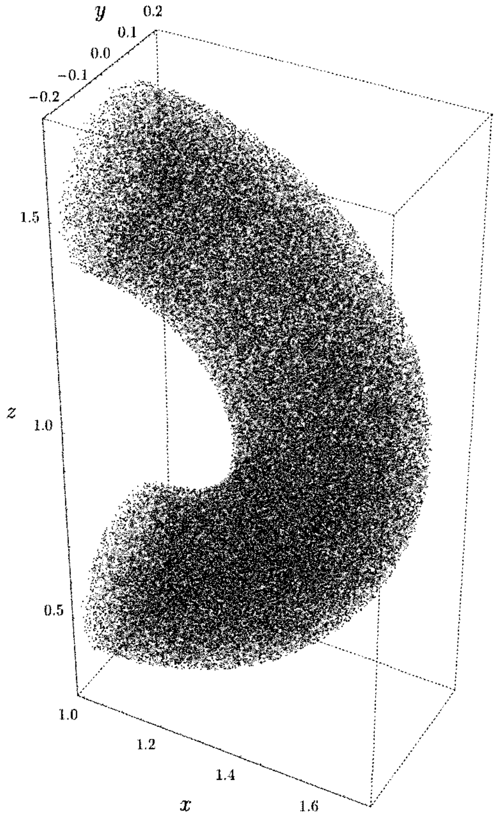
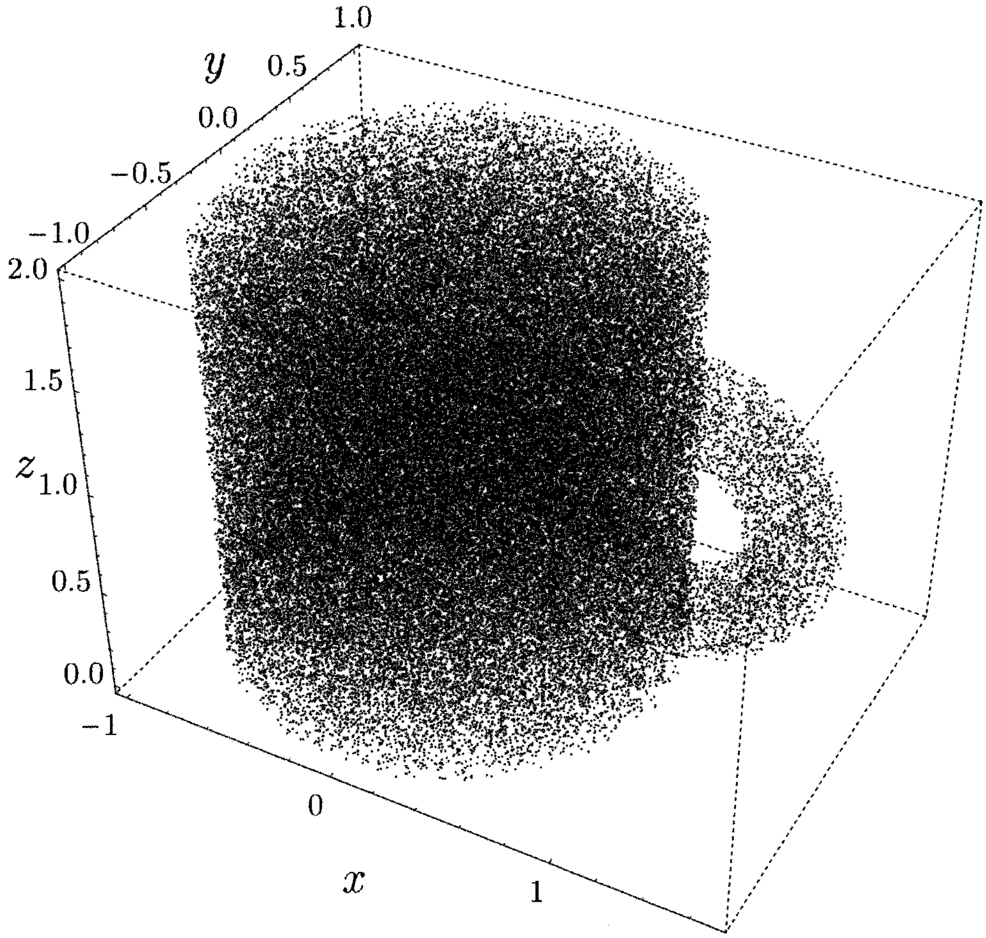


Figure 7: H distribution



The final step in generating a coffee mug distribution is to combine these two distributions such that the overall distribution is uniform. This can be achieved by considering the relative volumes of each region; the ratio of volumes of M and H is $r_M^2 h : \pi R^2 r_H$. Thus, it would be sufficient for $\frac{r_M^2 h}{r_M^2 h + \pi R^2 r_H}$ of points to be chosen from M, and the rest from H. For example, in 10^5 points chosen in S with the values in Figures 6 and 7, approximately 96954 points are chosen from M, and the remaining 3046 points are from H. This final result is illustrated in Figure 8.

Figure 8: A uniform coffee mug distribution



5.2 The generalized algorithm

The method to generate any desired region or boundary can be summarized in the following algorithm:

1. Begin with n-dimensional region R. Divide R into continuous subregions $\{R_1, R_2, ...\}$ such that union of all R_i is R. As this may be difficult to implement in practice, these subregions are best decided by human input.

never explained why.

2. For each R_i, compute using its parameterization a Jacobian determinant, multiplying the parametric equations by t_n if R_i is a boundary.

What a parametric equation is never explained

3. For each R_i, compute its total volume by integrating the Jacobian determinant over the complete region, obtaining α_i. Obtain expressions for the cumulative distribution functions $F_{R_i,T_j}(t_j)$ for each parameter t_j of each region R_i through integrating com- Never clearly arrived at

Why this works not explained.

not explained.

pletely the Jacobian determinant divided by α_i over all other parameters and partially for t_j, then calculate (exactly or numerically) their inverses.

even this never stated!

4. For each random point on R, choose an R_i randomly using weighted probabilities, with weights α_i. Then generate with a PRNG n uniformly distributed random values over $[0, 1)$ and transform these values by the inverses calculated for the R_i. Apply the parametric equations over the transformed values to obtain the desired Cartesian point.

6 Useful solutions

Section 5 presents a solution algorithm in which sections of complex regions are modularized into smaller, manageable subregions. This suggests that subregions that are common to many complex regions or a class of regions may appear often, and the inverse transformations for these subregions can be computed before even R is known. Examples of these subregions are simplices (especially for polytopes) and quadric surfaces. Table 1 lists some of these subregions, their parameterizations, and their required transformations $t_i(\psi_i)$. More details including the derivations of these solutions are included in Appendix B, in which Table 3 also illustrates the 'naive' uniform parameter approach and the corrected distributions after application of the following solutions.

Should be here

Table 1: Common subregions and their solutions

Region	Parameterization	Cumulative distributions	Solutions
n-simplex with vertices $\mathbf{p}_i$	$\mathbf{x}_0 = \mathbf{p}_0$; $\mathbf{x}_{n+1} = t_{n+1}x_n +$ $(1 - t_{n+1})\mathbf{p}_{n+1}$	$F_{T_i}(t_i) = t_i^i$	$t_i = \sqrt[i]{\psi_i}$?
Ellipsoid	$x = ar\cos\theta\sin\phi$ $y = br\sin\theta\sin\phi$ $z = cr\cos\phi$	$F_\Theta(\theta) = \theta/2\pi$ $F_\Phi(\phi) = \sin^2(\phi/2)$ $F_R(r) = r^3$	$\theta = 2\pi\psi_1$ $\phi = 2\arcsin\sqrt{\psi_2}$ $r = \sqrt[3]{\psi_3}$?

Continued on next page

Table 1 – continued from previous page

Region	Parameterization	Cumulative distributions	Solutions
Paraboloid bounded by $z \leq m$	$x = ahr\cos\theta$ $y = bhr\sin\theta$ $z = ch^2$	$F_H(h) = c^2h^4/m^2$ $F_\Theta(\theta) = \theta/2\pi$ $F_R(r) = r^2$	$h = \sqrt[4]{m^2\psi_1/c^2}$ $\theta = 2\pi\psi_2$ $r = \sqrt{\psi_3}$
Single-sheet hyperboloid	$x = ar\cos\theta\cosh m$ $y = br\sin\theta\cosh m$ $z = c\sinh m$	See Appendix B.2.3 $F_\Theta(\theta) = \theta/2\pi$ $F_R(r) = r^2$	Numerical $\theta = 2\pi\psi_2$ $r = \sqrt{\psi_3}$
Double-sheet hyperboloid	$x = ar\cos\theta\sinh m$ $y = br\sin\theta\sinh m$ $z = c\cosh m$	See Appendix B.2.4 $F_\Theta(\theta) = \theta/2\pi$ $F_R(r) = r^2$	Numerical $\theta = 2\pi\psi_2$ $r = \sqrt{\psi_3}$
Cone between $z_0 \leq z < z_1$	$x = arz\cos\theta$ $y = brz\sin\theta$ $z = z$	$F_Z(z) = \frac{z^3 - z_0^3}{z_1^3 - z_0^3}$ $F_\Theta(\theta) = \theta/2\pi$ $F_R(r) = r^2$	$z = (z_0^3 + (z_1^3 - z_0^3)\psi_1)^{1/3}$ $\theta = 2\pi\psi_2$ $r = \sqrt{\psi_3}$
Cylinder between $z_0 \leq z < z_1$	$x = ar\cos\theta$ $y = br\sin\theta$ $z = z$	$F_Z(z) = \frac{z - z_0}{z_1 - z_0}$ $F_\Theta(\theta) = \theta/2\pi$ $F_R(r) = r^2$	$z = z_0 + (z_1 - z_0)\psi_1$ $\theta = 2\pi\psi_2$ $r = \sqrt{\psi_3}$

?

None of this is justified. — some in Appendix B, but it would have been far better to explain the basics than to do all this.

7 Conclusion

7.1 Evaluation

My solution framework modularizes a single complex problem into smaller units. This process is greatly reminiscent of a much more common and simpler problem: volume calculation of nonregular solids. Consider the coffee cup example of Section 5; to calculate the volume, one would have to calculate the volumes of the half torus and the cylinder separately and add them to find the total volume. This seems almost identical to my method, which has simply been adapted to consider distributions rather than direct volumes, leading to some interesting implications. Nonreducible problems in volumes are also nonreducible in distributions. As my method involves a volume calculation step, distribution computation is strictly a 'harder' problem than volume computation. Consequently, the 'unit' solutions for distributions form strictly a subset of the set of 'unit' solutions for volumes.

Being 'harder' than volume calculation with almost the same method, uniform distribution computation inherits various problems. Some regions cannot be easily broken down precisely into convenient subregions. For example, the coffee cup was approximated as a cylinder with a half torus; however, as the cylinder has a curved surface which is connected to the flat cross section of the half torus, there is a small region between the cylinder and the half torus that is within the coffee cup but not considered in the final distribution. While for practical purposes this imperfection may be quite unnoticeable, mathematically the two distributions are not equivalent. Sometimes a complex region is also nonreducible, in which case the volume and in effect the distribution must be entirely numerically complicated for nonintegratable parameterizations.

Evidently, a more complex method also brings additional problems. For example, the half torus component of the distribution could not be completely symbolically computed, involving a numerical inverse. Not only is this mathematically a problem, but methods used to compute numerical inverses can often lead to roundoff errors or other inaccuracies (Abate, Choudhury, and Whitt 15). The existence of the inverse in the first place can be a concern, with the half torus being limited by $3r_H \geq 2R$. Ultimately, this can only be solved by choosing 'good' parameterizations for each unique problem.

7.2 Further investigation

Although my method provides solutions in the general case, these solutions are by no means the only possible ways to generate uniform distributions over arbitrary regions. For example, to generate a unit sphere, it may be sufficient to simply generate points in the unit cube and eliminate points outside the sphere. Alternatively, methods such as generating points across the sphere 'naively' as shown in Section 2.1, then removing points around the denser center with a normalizing probability as per the Jacobian, eliminate the need to use the computationally heavy inverse sine function. So, further analysis and benchmarking of different sampling techniques may help elucidate the differences between these methods.

Clearly, distributions over regions are also not limited to being uniform. However, how to generalize common one-dimensional distributions to higher dimensions is not entirely obvious. Despite this, some work has been done in this direction, and the normal distribution has been generalized to the multivariate case to produce Gaussian distributions over a plane (Do). In any case, it is difficult to reconcile the concept of these random variables, distributed over infinite spatial regions, to discrete, finite regions of space. Additionally, while these solutions are quite useful for cases up to three dimensions, where Monte Carlo modeling methods become commonly used approximations of reality, the practical application of additional dimensions that my solution offers is still unclear. Even so, an examination of the uniform case may provide insight into how to formalize these generalizations for future investigation.

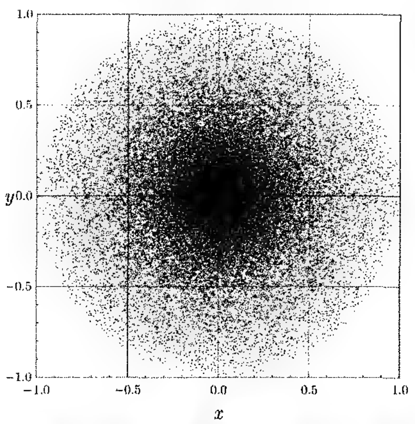
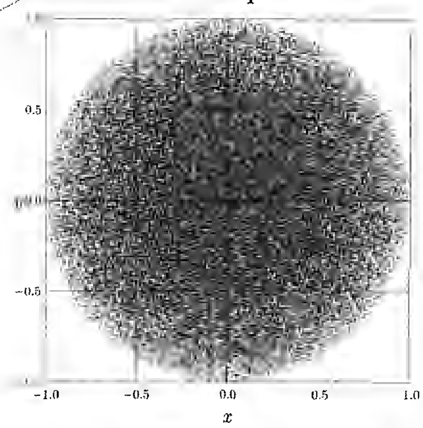
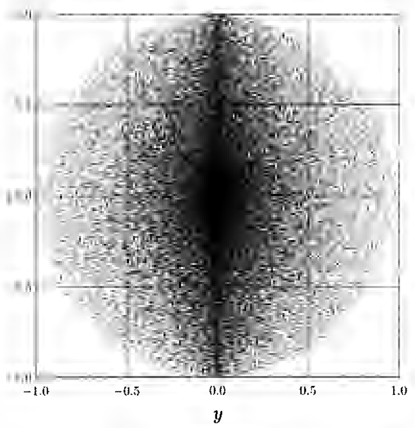
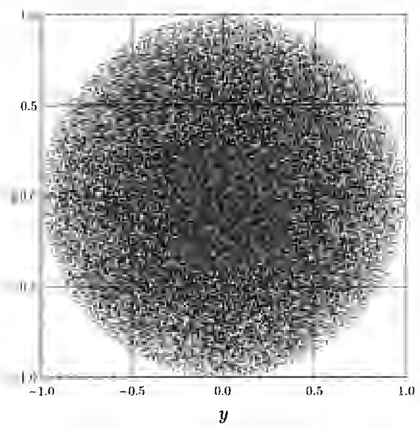
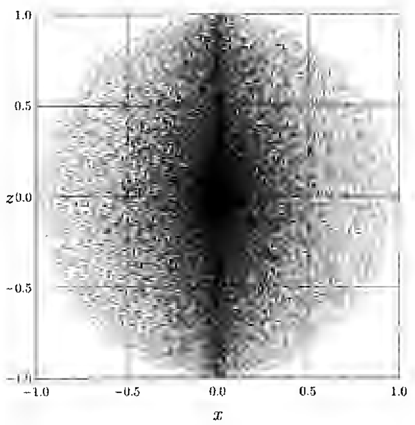
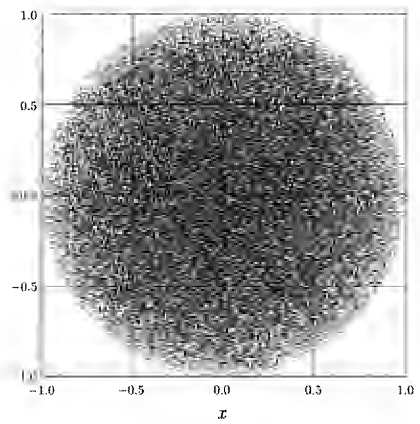
Hugely flawed - which the supervisor ought to have picked up.
The starting point for an EE is the HL syllabus & any Maths beyond this needs to be clearly explained an justified.
This was not the case here, so clear understanding was not shown.

Appendix A Additional figures

This shouldn't be here. It should be on page 3* before the sphere & should be fully explained

* and all on p. 6.

Table 2: Projections of nonuniform and uniform sphere distributions onto axis planes

Projection	Nonuniform sphere	Uniform sphere
xy-plane		
yz-plane		
xz-plane		

How this was produced is not explained at all.

Figure 9: Projections of coffee mug M distribution onto axis planes

xy-plane | yz-plane | xz-plane

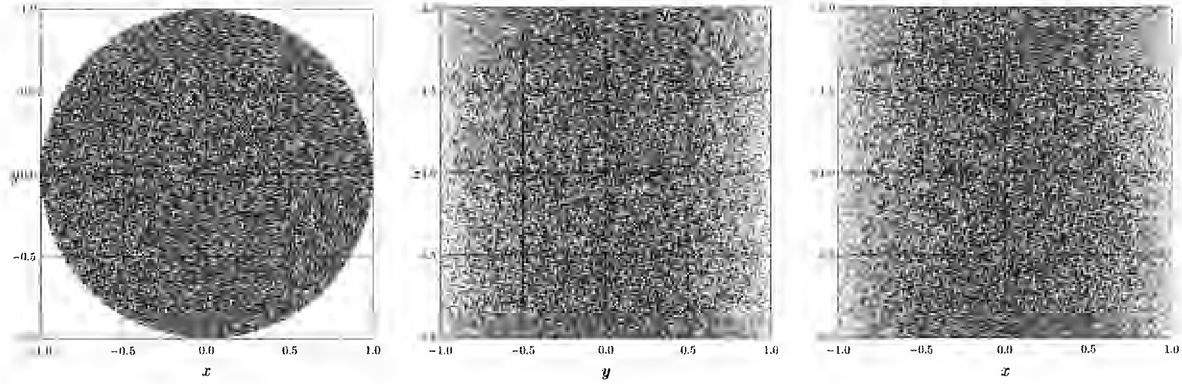
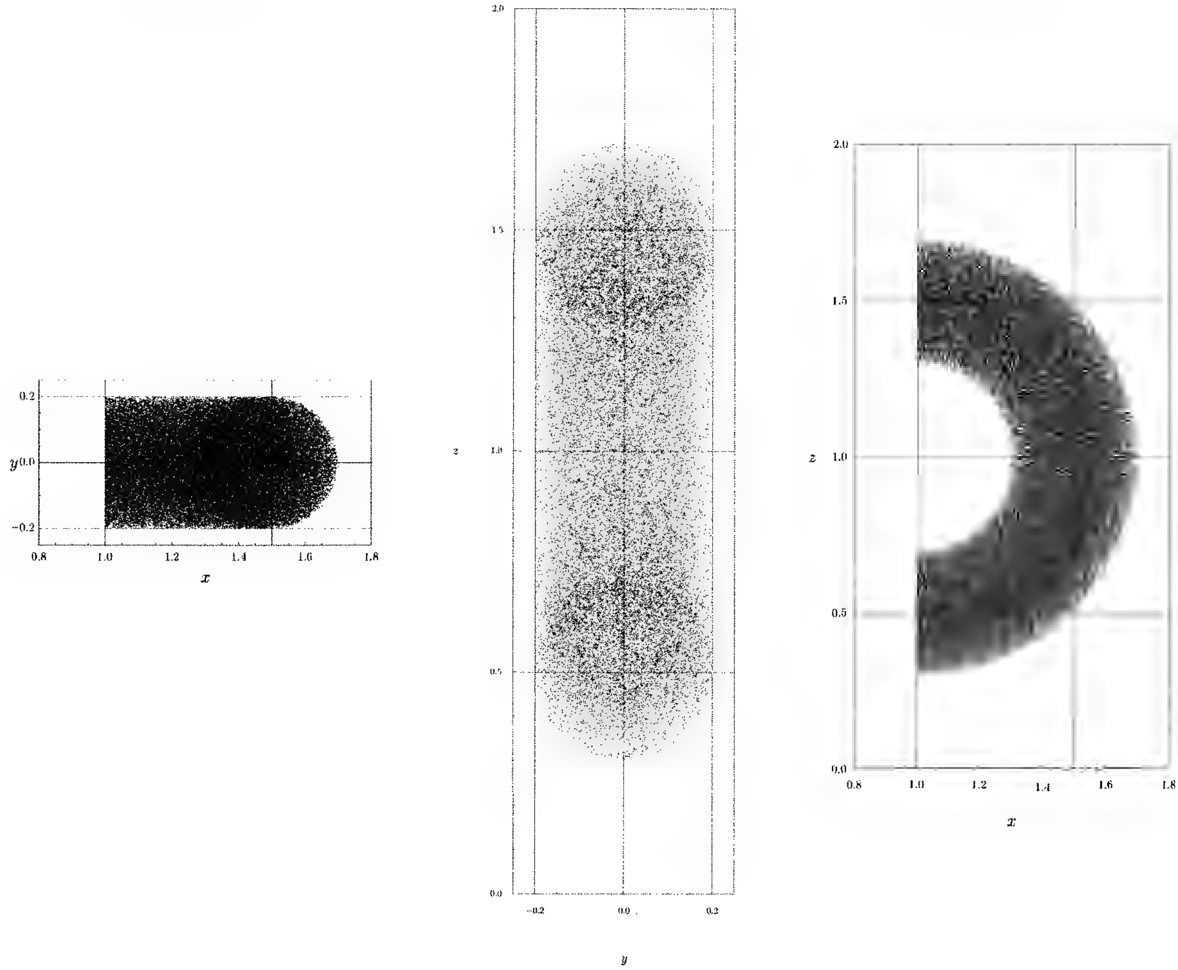


Figure 10: Projections of coffee mug H distribution onto axis planes

xy-plane | yz-plane | xz-plane



Appendix B Derivations and illustrations of Section 6 solutions

B.1 n-simplical regions

Let the vertices of the n-simplex be column vectors $\mathbf{p}_0, \mathbf{p}_1, ..., \mathbf{p}_n$. The 0-simplex has parameterization $\mathbf{x}_0() = \mathbf{p}_0$. The 1-simplex (line segment) has the well-known parameterization $\mathbf{x}_1(t_1) = t_1\mathbf{p}_0 + (1 - t_1)\mathbf{p}_1$. $(n+1)$-simplices can be formed with an infinite number of line segments radiating from the next vertex, $\mathbf{p}_{n+1}$, to all points on a corresponding n-simplex. Hence, the n-simplex parameterization follows the recurrence relation $\mathbf{x}_{n+1}(t_1, t_2, ..., t_{n+1}) = t_{n+1}\mathbf{x}_n(t_1, t_2, ..., t_n) + (1 - t_{n+1})\mathbf{p}_{n+1}$.

For the Jacobian, we must compute each column vector $\partial\mathbf{x}_n/\partial t_i$. The recurrence relation implies $\partial\mathbf{x}_n/\partial t_n = \mathbf{x}_{n-1} - \mathbf{p}_n$. For $i < n$, $\partial\mathbf{x}_n/\partial t_i = t_n\, \partial\mathbf{x}_{n-1}/\partial t_i$, so with reductions of the RHS until $\partial\mathbf{x}_i/\partial t_i$, we can obtain:

$$\frac{\partial\mathbf{x}_n}{\partial t_i} = t_n t_{n-1} \cdots t_{i+1}\frac{\partial\mathbf{x}_i}{\partial t_i} = t_{i+1}t_{i+2}\cdots t_n(\mathbf{x}_{i-1} - \mathbf{p}_i)$$

The Jacobian determinant of the region is therefore:

$$\begin{aligned}\left|\frac{\partial\mathbf{x}_n}{\partial\mathbf{t}}\right| &= \left|\frac{\partial\mathbf{x}_n}{\partial t_1} \quad \frac{\partial\mathbf{x}_n}{\partial t_2} \quad \cdots \quad \frac{\partial\mathbf{x}_n}{\partial t_n}\right| \\ &= \left|t_2t_3\cdots t_n(\mathbf{x}_0 - \mathbf{p}_1) \quad t_3t_4\cdots t_n(\mathbf{x}_1 - \mathbf{p}_2) \quad \cdots \quad \mathbf{x}_{n-1} - \mathbf{p}_n\right| \\ &= c(t_2t_3t_4\cdots t_n)(t_3t_4\cdots t_n)\cdots(t_n) \\ &= ct_2^1t_3^2t_4^3\cdots t_n^{n-1}\end{aligned}$$

Here, $c = |\mathbf{x}_0 - \mathbf{p}_1 \quad \mathbf{x}_1 - \mathbf{p}_2 \quad \cdots \quad \mathbf{x}_{n-1} - \mathbf{p}_n|$. Consider the recurrence relation rearrangement $\mathbf{x}_i - t_i(\mathbf{x}_{i-1} - \mathbf{p}_i) = \mathbf{p}_i$ and that matrix determinants remain constant under column addition. Hence, by subtracting each kth column by the $(k-1)$th column with factor t_k for $k \geq 2$:

$$c = |\mathbf{p}_0 - \mathbf{p}_1 \quad \mathbf{p}_1 - \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_{n-1} - \mathbf{p}_n|$$

Thus, c is independent of the parameters t_i. Finally:

$$\begin{aligned}\alpha &= \int_0^1 \cdots \int_0^1 ct_2^1t_3^2t_4^3\cdots t_n^{n-1}\, dt_1 ... dt_n \\ &= \frac{c}{n!}\end{aligned}$$

$$
\begin{aligned}
F_{T_i}(t_i) &= n! \int_0^{t_i} \int_0^1 \cdots \int_0^1 t_2^1 t_3^2 t_4^3 \cdots t_n^{n-1} \, dt_1 \ldots dt_{i-1} \, dt_{i+1} \ldots dt_n \, dt_i \\
&= t_i^i \\
t_i(\psi_i) &= \sqrt[i]{\psi_i}
\end{aligned}
$$

B.2 Quadric surfaces

B.2.1 Ellipsoid

Consider an ellipsoid satisfying $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$. The region has the well-known parameterization $\mathbf{x}(\theta, \phi, r) = (ar\cos\theta\sin\phi, br\sin\theta\sin\phi, cr\cos\phi)$ for $0 \le \theta < 2\pi$, $0 \le \phi < \pi$, and $0 \le r < 1$. $|\partial\mathbf{x}/\partial\mathbf{t}| = -abcr^2 \sin\phi$ and $\alpha = -4\pi abc/3$:

$$
\begin{aligned}
F_\Theta(\theta) &= \frac{\theta}{2\pi} &\rightarrow\quad& \theta(\psi_1) = 2\pi\psi_1; \\
F_\Phi(\phi) &= \sin^2 \frac{\phi}{2} &\rightarrow\quad& \phi(\psi_2) = 2 \arcsin \sqrt{\psi_2}; \\
F_R(r) &= r^3 &\rightarrow\quad& r(\psi_3) = \sqrt[3]{\psi_3}
\end{aligned}
$$

B.2.2 Paraboloid

Consider the region bounded by the paraboloid $x^2/a^2 + y^2/b^2 - z/c = 0$ and the plane $z = m$. This region can be constructed as an infinite number of ellipses parallel to the z-axis between the origin and $z = m$. Hence, with $z = ch^2$, where $0 \le h < \sqrt{m/c}$, the elliptical regions can be produced for each z as $x = ahr\cos\theta$ and $y = bhr\sin\theta$, where $0 \le r < 1$ and $0 \le \theta < 2\pi$ ($r = 1$ produces the paraboloid itself). Note that this is not a typical region that is generated by multiplying by some t_3. $|\partial(x, y, z)/\partial(h, \theta, r)| = -2abch^3 r$ and $\alpha = -\pi abm^2/2c$:

$$
\begin{aligned}
F_H(h) &= \frac{c^2 h^4}{m^2} &\rightarrow\quad& h(\psi_1) = \sqrt[4]{\frac{m^2\psi_1}{c^2}}; \\
F_\Theta(\theta) &= \frac{\theta}{2\pi} &\rightarrow\quad& \theta(\psi_2) = 2\pi\psi_2; \\
F_R(r) &= r^2 &\rightarrow\quad& r(\psi_3) = \sqrt{\psi_3}
\end{aligned}
$$

B.2.3 Single-sheet hyperboloid

Consider the region bounded by the hyperboloid $x^2/a^2 + y^2/b^2 - z^2/c^2 = 1$ and the planes $z = c\sinh m_0$ and $z = c\sinh m_1$ (with $m_1 > m_0$). Like the paraboloid, this may be parametrized using an infinite number of ellipses; using this, an appropriate parameterization is $\mathbf{x}(m, \theta, r) = (ar\cos\theta\cosh m, br\sin\theta\cosh m, c\sinh m)$ for $m_0 \le m < m_1$. Hence, the Jacobian $|\partial \mathbf{x}/\partial \mathbf{t}| = -abcr\sinh^3 m$ and $\alpha = -\pi abc\,(g(m_1) - g(m_0))/12$, where $g(m) = 9\sinh m + \sinh 3m$:

$$\begin{aligned} F_M(m) &= \frac{g(m) - g(m_0)}{g(m_1) - g(m_0)}; \\ F_\Theta(\theta) &= \frac{\theta}{2\pi} \quad \rightarrow \quad \theta(\psi_2) = 2\pi\psi_2; \\ F_R(r) &= r^2 \quad \rightarrow \quad r(\psi_3) = \sqrt{\psi_3} \end{aligned}$$

B.2.4 Double-sheet hyperboloid

Consider the region bounded by the hyperboloid $x^2/a^2 + y^2/b^2 - z^2/c^2 = -1$ and the plane $z = c\cosh m_1$. Like the single-sheet hyperboloid, this may be parametrized using an infinite number of ellipses beginning from $z = c$; an appropriate parameterization is $\mathbf{x}(m, \theta, r) = (ar\cos\theta\sinh m, br\sin\theta\sinh m, c\cosh m)$ for $0 \leq m < m_1$. $|\partial \mathbf{x}/\partial \mathbf{t}| = -abcr\cosh^3 m$ and $\alpha = -\pi abc\,(g(m_1) - 8)/12$, where $g(m) = 9\cosh m - \cosh 3m$:

$$\begin{aligned} F_M(m) &= \frac{g(m) - 8}{g(m_1) - 8}; \\ F_\Theta(\theta) &= \frac{\theta}{2\pi} \quad \rightarrow \quad \theta(\psi_2) = 2\pi\psi_2; \\ F_R(r) &= r^2 \quad \rightarrow \quad r(\psi_3) = \sqrt{\psi_3} \end{aligned}$$

B.2.5 Cone

Consider the region bounded by the cone $x^2/a^2 + y^2/b^2 - z^2 = 0$ and the planes $z = z_0$ and $z = z_1$ (with $z_0 < z_1$). A common parameterization is $\mathbf{x}(z, \theta, r) = (arz\cos\theta, brz\sin\theta, z)$ for $z_0 \leq z < z_1$. $|\partial \mathbf{x}/\partial \mathbf{t}| = -abrz^2$ and $\alpha = -\pi ab\,(z_1^3 - z_0^3)/3$:

$$\begin{aligned} F_Z(z) &= \frac{z^3 - z_0^3}{z_1^3 - z_0^3} \quad \rightarrow \quad z(\psi_1) = \sqrt[3]{(z_1^3 - z_0^3)\psi_1 + z_0^3}; \\ F_\Theta(\theta) &= \frac{\theta}{2\pi} \quad \rightarrow \quad \theta(\psi_2) = 2\pi\psi_2; \\ F_R(r) &= r^2 \quad \rightarrow \quad r(\psi_3) = \sqrt{\psi_3} \end{aligned}$$

B.2.6 Elliptic cylinder

Consider the region bounded by the cylinder $x^2/a^2 + y^2/b^2 = 1$ and the planes $z = z_0$ and $z = z_1$ (with $z_0 < z_1$). A common parameterization is $\mathbf{x}(z, \theta, r) = (ar\cos\theta, br\sin\theta, z)$ for

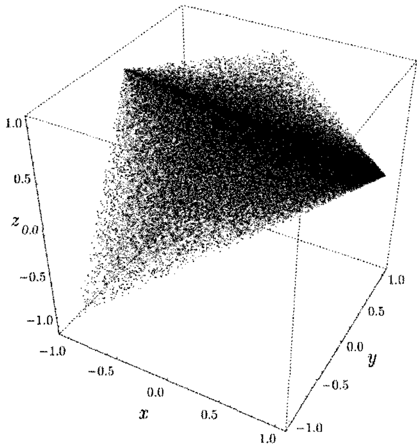
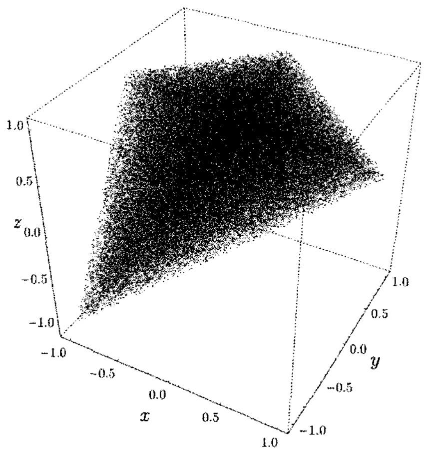
$z_0 \leq z < z_1$. $|\partial \mathbf{x}/\partial \mathbf{t}| = -abr$ and $\alpha = -\pi ab\,(z_1 - z_0)$:

$$\begin{aligned}
F_Z(z) &= \frac{z - z_0}{z_1 - z_0} &\quad\rightarrow\quad& z(\psi_1) = (z_1 - z_0)\psi_1 + z_0; \\
F_\Theta(\theta) &= \frac{\theta}{2\pi} &\quad\rightarrow\quad& \theta(\psi_2) = 2\pi\psi_2; \\
F_R(r) &= r^2 &\quad\rightarrow\quad& r(\psi_3) = \sqrt{\psi_3}
\end{aligned}$$

B.3 Illustrations of subsection solutions

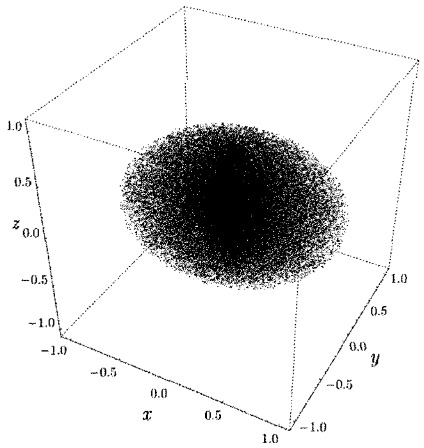
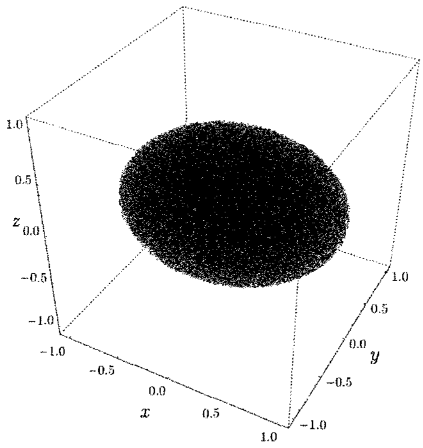
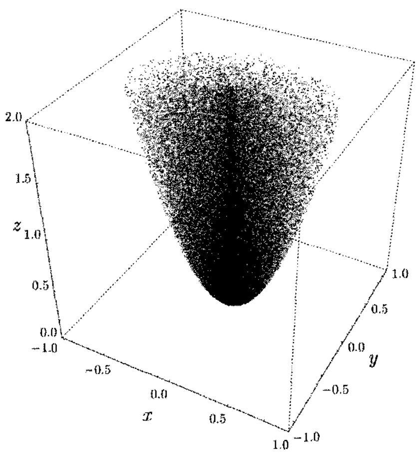
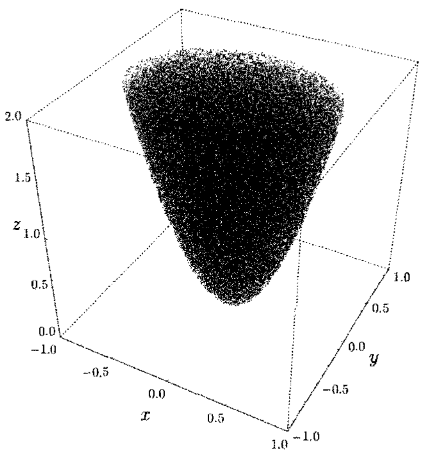
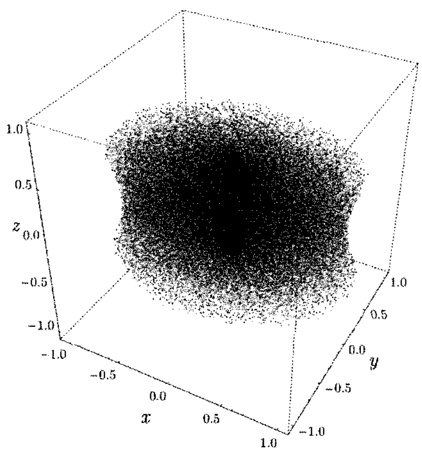
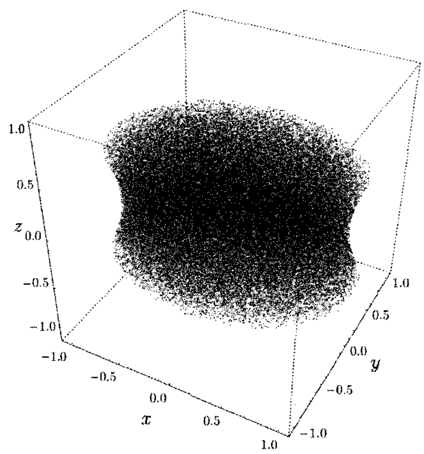
For the variable geometric specifications for each region, refer to the corresponding parameterization of the region. All of the corrected distributions were generated using the solutions presented in Table 1 with 10^5 points each.

Table 3: Illustrations of 'naive' and corrected distributions over Section 6 subregions

Region and specifications	Uniform parameter distribution	Corrected distribution
3-simplex $\mathbf{p}_0 = (-1,-1,-1)$ $\mathbf{p}_1 = (0,1,1)$ $\mathbf{p}_2 = (-1,0,1)$ $\mathbf{p}_3 = (1,1,0)$		

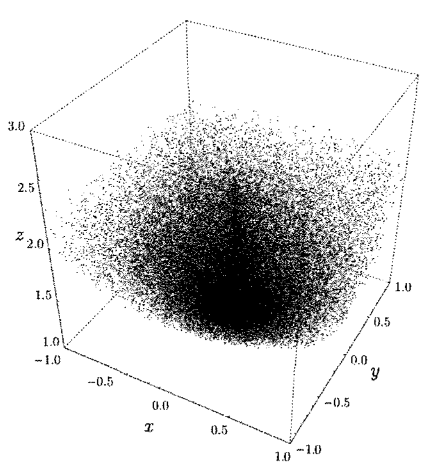
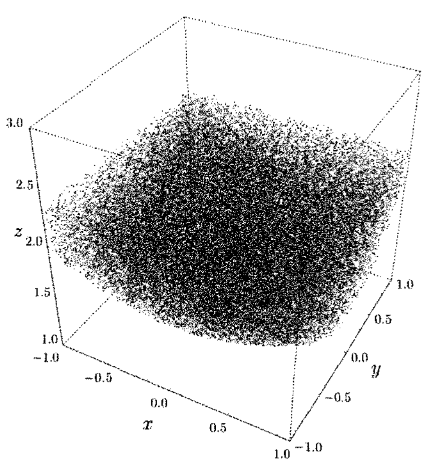
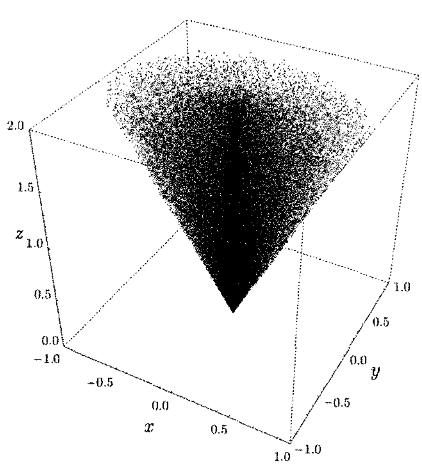
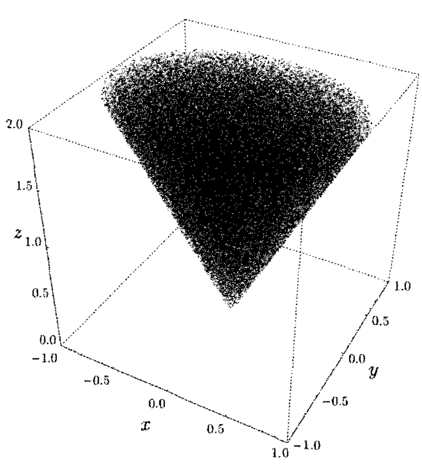
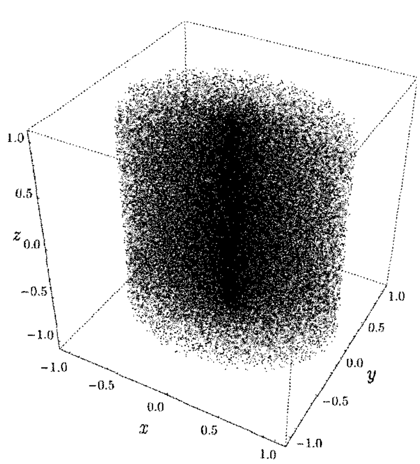
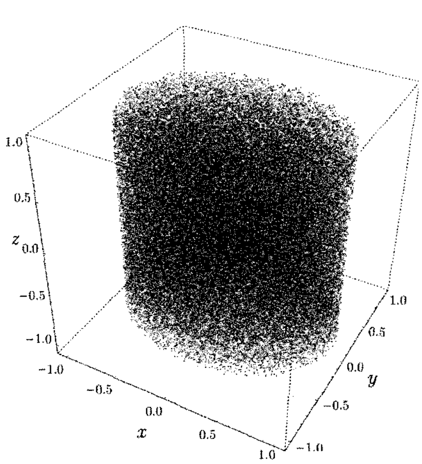
Continued on next page

Table 3 – continued from previous page

Region and specifications	Uniform parameter distribution	Corrected distribution
Ellipsoid $a = 1$ $b = 0.8$ $c = 0.6$		
Paraboloid $m = 2$ $a = 1$ $b = 0.75$ $c = 3$		
Single-sheet hyperboloid $m_0 = -0.5$ $m_1 = 0.5$ $a = 1$ $b = 0.75$ $c = 1$		

Continued on next page

Table 3 – continued from previous page

Region and specifications	Uniform parameter distribution	Corrected distribution
Double-sheet hyperboloid $m_1 = 1.5$ $a = 1$ $b = 0.75$ $c = 1$		
Cone $z_0 = 0$ $z_1 = 2$ $a = 1$ $b = 0.75$		
Cylinder $z_0 = -1$ $z_1 = 1$ $a = 1$ $b = 0.75$		

Works Cited

Abate, Joseph, Gagan L. Choudhury, and Ward Whitt. "An Introduction to Numerical Transform Inversion and Its Application to Probability Models." *Computational Probability.* Ed. W. Grassman. Boston: Kluwer, 1999. 257-323. Web.

Amidror, Isaac. "The Jacobian Determinant of a 2D Transformation and Its Significance." *The Theory of the Moiré Phenomenon: Volume II: Aperiodic Layers.* Dordrecht: Springer, 2007. 309-26. Print.

Billingsley, Patrick. *Probability and Measure.* 3rd ed. New York: Wiley, 1995. Print.

Do, Chuong B. *The Multivariate Normal Distribution.* N.p.: Stanford University, 8 Oct. 2008. PDF.

Gelman, Andrew. "Prior Distributions for Variance Parameters in Hierarchical Models." *Bayesian Analysis* 1 (2006): 515-33. Web.

Penzov, A. A., I. T. Dimov, N. M. Mitev, G. I. Shirakova, and L. Szirmay-Kalov. "Examining the Distribution of Sampling Point Sets on Sphere for Monte Carlo Image Rendering." *AIP Conference Proceedings* (2010): n. pag. *Examining the Distribution of Sampling Point Sets on Sphere for Monte Carlo Image Rendering.* 1 June 2014. Web. 15 Oct. 2014.