Monte Carlo II - Random Sampling

Lecture #14: Thursday 13 March 2003
Lecturer: Greg Humphreys
Scribe: John Tran
Reviewer: Amanda Huganiss

This is the second lecture on Monte Carlo techniques. This lecture focuses on random sampling strategies.

1 Probability Distribution Functions

How do we draw samples from some arbitrary PDF $p(x)$? Assume we have a steady supply of uniform random variables $U_1, U_2, \ldots, U_N$ (for example using drand48()). Furthermore, assume these random variables are between zero and one.

1.1 Discrete Probability Distributions

For a PDF as shown in the left figure in figure **??**, construct a cumulative distribution function as in the right figure in figure **??**.

For the PDF, $p_i > 0$ and $\sum_i^N p_i = 1$. For the CDF, $P_j = \sum_{i=1}^j p_i$.

Notice that if we project the CDF onto the Y axis, the heights of the intervals recover p_i.

Now select event X_i if $P_{i-1} < U \leq P_i$! Think of this as projecting a uniform variable onto the Y axis, then projecting it down.

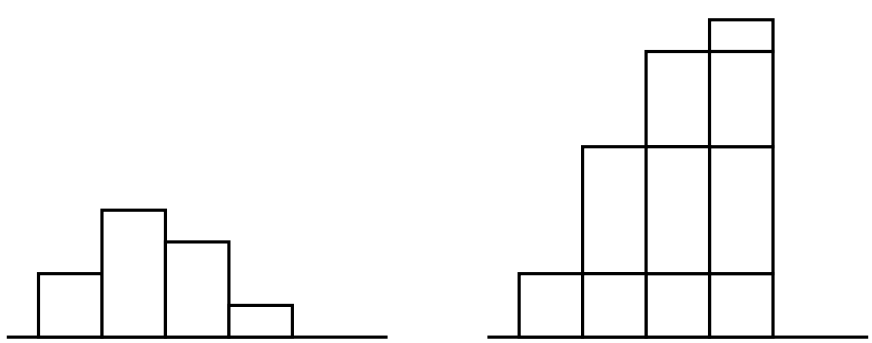

Figure 1: A PDF (left) and its CDF (right)

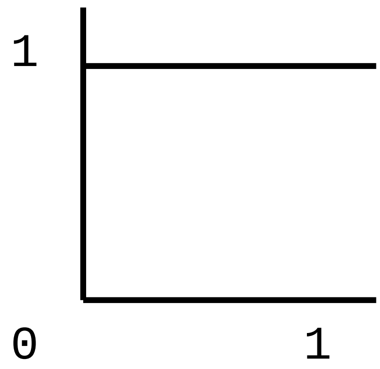

Figure 2: A uniform PDF

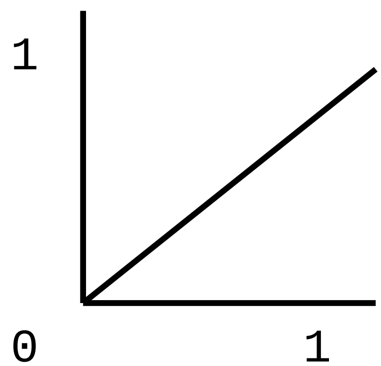

Figure 3: A uniform CDF

1.2 Continuous Probability Distributions

Recall for a PDF, $p(x) \geq 0$, and $\int p(x)dx = 1$. The CDF is exactly like our P_i from before, such that $P(x) = \int_0^x p(t)dt$.

1.3 Uniform PDF

Given a uniform variable, the PDF looks like figure **??**.

The uniform CDF then looks like figure **??**.

This means that $P(x)$ is the probability that $X < x$, with $P(1) = 1$.

1.4 Gaussian Distribution

Given a Gaussian distribution, the PDF looks like figure **??**

The uniform CDF then looks like figure **??**.

To sample from this $p(x)$, project U from the Y axis, then down to get X.

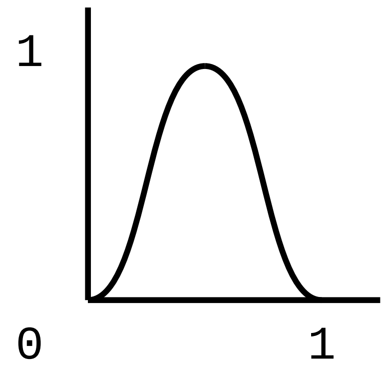

Figure 4: A Gaussian PDF

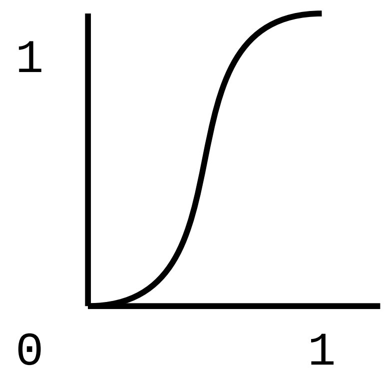

Figure 5: A Gaussian CDF

What are we doing here? This is simply function inversion!

1.5 Sampling from any distribution function

The steps to sample from any distribution are:

1. Compute the CDF, $P(x) = \int_0^x p(t)dt$
2. Invert P(x)
3. Sample from $X = P^{-1}(U)$

1.6 Example: Power Distribution

Let $p(x) \propto x^n$. We know $\int_0^1 p(x)dx = 1$, so $\int_0^1 cx^n dx = 1$. We first have to determine the normalization factor, c.

$$\left.\frac{cx^{n+1}}{n+1}\right|_0^1 = 1$$

$$\frac{c}{n+1} = 1$$

$$c = n+1$$

Therefore:

$$p(x) = (n+1)x^n$$

Integrating this, we get:

$$P(x) = x^{n+1}$$

Inverting this, we get:

$$P^{-1}(x) = \sqrt[n+1]{x}$$

So, finally, we get our random variable:

$$X = \sqrt[n+1]{x} \tag{1}$$

1.6.1 A Trick for the Power Distribution

$Y = max(U_1, U_2, \ldots, U_{n+1})$ is a power distribution! This is because the probability that $\{Y < x\}$ is the probability that *ALL* the $U_i < x$, but the U_i are independent!, so

$$\Pr\{Y < x\} = \prod_{i=1}^{n+1} \underbrace{\Pr\{U_i < x\}}_{x} = x^{n+1} \tag{2}$$

This method is faster for small n (like 2).

1.6.2 A Trick for a Gaussian Distribution

Take the *average*! This is a consequence of the central limit theorem, and the proof is left as an exercise to the reader.

2 Importance sampling

Suppose we have the integral,

$$I = \int_0^\infty e^{-cx} g(x) dx \tag{3}$$

This integral comes up in volume rendering.

It is useful to sample from a PDF that is *similar* to the integrand. This is called *importance sampling*. Looking at Equation **??**, we could guess at a PDF of the form

$$p(x) \propto e^{-cx} \tag{4}$$

First, we normalize:

$$\int_0^\infty ke^{-cx} dx = k\left(-\frac{1}{c}\right)\left(e^{-cx}\right)\Big|_0^\infty \tag{5}$$

$$\int_0^\infty ke^{-cx} dx = -\frac{k}{c}(0-1) \tag{6}$$

$$\int_0^\infty ke^{-cx} dx = \frac{k}{c} \tag{7}$$

$$\int_0^\infty ke^{-cx} dx = 1 \Rightarrow k = c \tag{8}$$

So we have $p(x) = ce^{-cx}$. Now we integrate this!

$$P(x) = \int_0^x ce^{-ct} dt \tag{9}$$

$$P(x) = -\frac{c}{c}\left[e^{-ct}\right]_0^x \tag{10}$$

$$P(x) = -\left(e^{-cx} - 1\right) \tag{11}$$

$$P(x) = 1 - e^{-cx} \tag{12}$$

Now we invert the function.

$$U = 1 - e^{-cx} \tag{13}$$

$$e^{-cx} = 1 - U \tag{14}$$

$$-cx = \ln(1 - U) \tag{15}$$

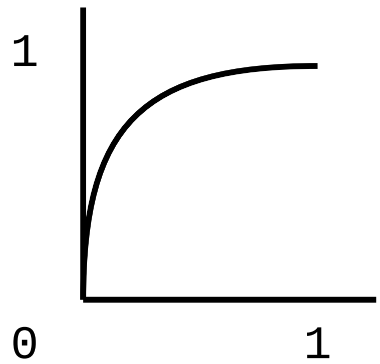

Figure 6: $P(x)$ for I

$$x = -\frac{-\ln(1-U)}{c} \tag{16}$$

If U is uniform, then so is $1-U$, so this becomes:

$$X = -\frac{\ln U}{c} \tag{17}$$

Note that $P(x)$ looks like Figure **??**.

This means that there is a high probability of sampling small X. This makes sense because ce^{-cx} goes to zero quickly if x is big, so big x contributes little to the integral.

Remember, applying Monte Carlo is easy:

$$F_N = \frac{1}{N}\sum_{i=1}^{N}\frac{e^{-cx_i}g(x_i)}{ce^{-cx_i}} \tag{18}$$

$$F_N = \frac{1}{cN}\sum_{i=1}^{N}g(x_i) \tag{19}$$

There are no exponentials! This is a much better sampling strategy and it is easy to apply. Simply draw X_i as $-\frac{1}{c}\ln U$.

3 Generalization: Transformation of Variables

The inversion method requires uniform X_i. Suppose we have X_i drawn from some arbitrary PDF $p_x(x)$. Now let $Y = y(X)$. What is the PDF of Y?

Assume $y(x)$ is continuous and non-decreasing (i.e. $y'(x) \geq 0$). We know that

$$\Pr\{Y \leq y(x)\} = \Pr\{X \leq x\} \tag{20}$$

so

$$P_y(y) = P_x(x) \tag{21}$$

Remember that P is the CDF, so we get the PDF by differentiation

$$p_y(y)\frac{dy}{dx} = p_x(x) \tag{22}$$

or

$$p_y(y) = \left(\frac{dy}{dx}\right)^{-1} p_x(x) \tag{23}$$

If $y(x)$ can be non-increasing, this works out to

$$P_y(y) = \left(\frac{dy}{dx}\right)^{-1} P_x(x) \tag{24}$$

How does this work? Suppose $p_x(x) = 2x$ over [0,1], and let $Y = sin(X)$. What is the PDF of Y?

$$\frac{dy}{dx} = \cos(x) \tag{25}$$

$$p_y(y) = |\cos x|^{-1} P_x(x) \tag{26}$$

$$p_y(y) = (\cos x)^{-1} 2x \tag{27}$$

$$p_y(y) = \frac{2\sin^{-1}(y)}{\cos(\sin^{-1}(y))} \tag{28}$$

$$p_y(y) = \frac{2\sin^{-1}(y)}{\sqrt{1-y^2}} \tag{29}$$

This is kind of backwards - usually we have some PDF we want to sample from, not some transformation. If we have X from $p_x(x)$ and we want Y from $p_y(y)$, what $y(x)$ should we use? All we need is $P_Y(Y(= P_X(X)$, so

$$y(x) = P_Y^{-1}(P_X(x)) \tag{30}$$

Notice this is clearly a generalization of the inversion method, since for uniform X, $P_x(x) = x$. The advantage of the transformation method in 1D is that we can start with non-uniform samples. In higher dimensions, we can use this to go from one domain to another. The advantage of the inversion method is that it preserves stratification.

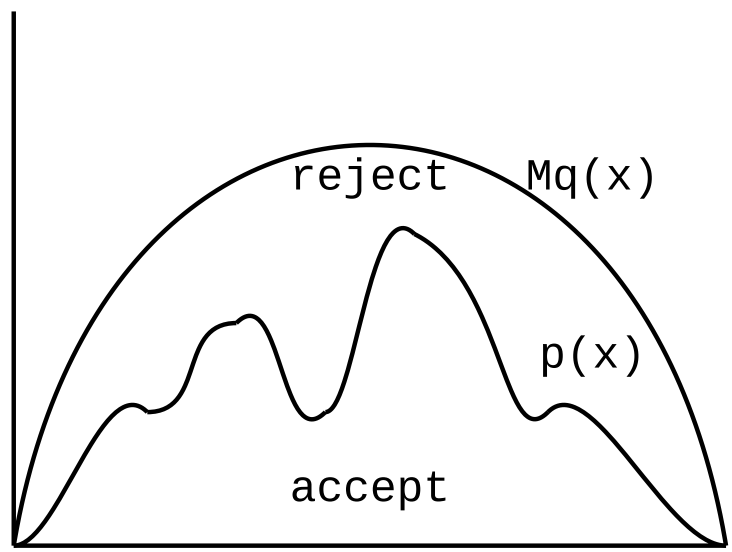

Figure 7: Rejection Sampling

4 Rejection Method

All this integrating is annoying. Let's just throw darts!

Let's say we want to sample from $p(x)$, but we don't know how to integrate it, or we don't know how to invert P_x). But there is another PDF $q(x)$ that satisfies

$$P(x) \leq Mq(x) \tag{31}$$

for some M, and we know how to sample from $q(x)$.

The reject algorithm is as follows:

reject:
  while (1)
    X = q→sample()
    U = uniform→sample()
    if U < $\frac{P(X)}{Mq(X)}$ return X

Basically, we choose two variables, X and U. If the point $(X, U \cdot Mq(X))$ lies under $p(x)$, accept point X.

Clearly, the efficiency of this technique depends on how tightly $Mq(x)$ bounds $p(x)$.

This works in any dimension! Let's do multidimensional sampling in the other order.

4.1 Rejection Sampling a Unit Circle

Generate a random (X, Y) in the square, and accept if (X, Y) is in the circle.

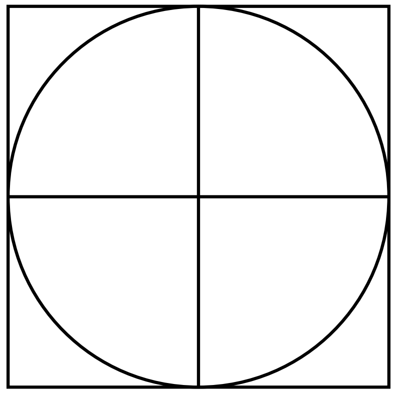

Figure 8: Rejection Sampling a unit circle

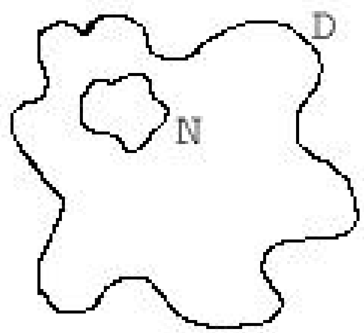

Figure 9: Probability in this case is just the ratio of the areas

do{
  X = 1 - 2U_1
  Y = 1 - 2U_2
} while ($X^2 + Y^2 > 1$)

This works for more complex shapes, as long as they have an inside-outside test.

Notice what we are doing. What does it mean to say that a 2D point is sampled uniformly? Uniform with respect to what? Uniform with respect to *area*! This means that if we pick *any* neighborhood inside the shape, the probability of choosing a point in that neighborhood is the area of the neighborhood divided by the area of the domain.

$$\Pr\{P \in N\} = \frac{A(N)}{A(D)} \tag{32}$$

5 Transformation Method in Multiple Dimensions

We have already seen transformation in 1D. Let's figure out how multidimensional PDF's are affected by transformation.

Suppose we have an n-D random variable $X = (X_1, X_2, \ldots, X_n)$ over some domain $\Omega_x \subset R^n$ with density function $p_x(x)$. Now let $Y = T(x)$, where $T : \Omega_x \longrightarrow \Omega_y$. (Note that T is a bijection).

What is the density function of Y? Because Y is a bijection, we know:

$$\Pr\{Y \in T(D)\} = \Pr\{X \in D\} \tag{33}$$

for *any* domain $D \subseteq \Omega_x$. We can write this in terms of the density functions p_x and p_y:

$$\int_{T(D)} P_Y(y)dy = \int_D P_X(x)dx \tag{34}$$

Now, apply a change of variables to the left hand side. This requires the determinant of the Jacobian:

$$\int_X P_Y(T(x))\,|J_T(x)|\,dx = \int_D P_X(x)dx \tag{35}$$

Recall:

$$J_T = \begin{bmatrix} \frac{\delta y_1}{\delta x_1} & \cdots & \frac{\delta y_1}{\delta x_n} \\ \vdots & \ddots & \vdots \\ \frac{\delta y_n}{\delta x_1} & \cdots & \frac{\delta y_n}{\delta x_n} \end{bmatrix} \tag{36}$$

But this equation is true for *any* domain D! So the integrands must be equal.

$$P_Y(T(x))\,|J_T(x)| = P_X(x) \tag{37}$$

or

$$P_Y(T(x)) = \frac{P_X(x)}{|J_T(x)|} \tag{38}$$

5.1 Example: Polar Coordinates

$$x = r\cos\theta$$

$$y = r\sin\theta$$

Suppose we draw samples from some density $p(r, \theta)$. What is $p(x, y)$? The Jacobian is

$$\begin{bmatrix} \frac{\delta x}{\delta r} & \frac{\delta x}{\delta \theta} \\ \frac{\delta y}{\delta r} & \frac{\delta y}{\delta \theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} \tag{39}$$

and the determinant is $r(\cos^2\theta + \sin^2\theta) = r$. So

$$p(x, y) = \frac{p(r, \theta)}{r} \tag{40}$$

This is kind of backwards. Usually we have $p(x.y)$ and we want $p(r, \theta)$, so

$$p(r, \theta) = rp(x, y) \tag{41}$$

5.2 Spherical Coordinates

$$x = r\sin\theta\cos\phi$$

$$y = r\sin\theta\sin\phi$$

$$z = r\cos\theta$$

The Jacobian has determinant $|J_T| = r^2\sin\theta$, so $p(r, \theta, \phi) = r^2\sin\theta p(x, y, z)$.

We can represent points on the surface of a unit sphere by setting $r = 1$. The reason this is important is that we often represent direction as a unit vector (x, y, z). Remember that solid angle is the area of a set of points on the unit sphere. In spherical coordinates, we get

$$d\omega = \sin\theta d\theta d\phi \tag{42}$$

So if we have a density function over solid angle, this just means that

$$\Pr\{\Omega \in D\} = \int_D p(\omega)d\omega \tag{43}$$

The density with respect to (θ, ϕ), is

$$p(\theta, \phi)d\theta\phi = p(\omega) \Rightarrow p(\theta, \phi) = \sin\theta d\omega \tag{44}$$

6 Sampling by Transformation in 2D

Suppose we have a joint density function $p(x, y)$. We want to draw samples from this distribution, but how? Use Conditional Probability!

Definition Marginal Density Function $p(x)$

$$p(x) = \int p(x, y)dy \tag{45}$$

This is just the density function for X alone.

Definition Conditional Density Function $p(y|x)$

$$p(y|x) = \frac{p(x,y)}{p(x)} \tag{46}$$

This gives the density function for Y once we have chosen a particular X.

The basic idea is that we want to compute the marginal density to isolate and sample one variables. Then, we use the conditional density to sample the other variable.

6.1 Example: Sample a Hemisphere

Suppose we want to sample a hemisphere uniformly with respect to solid angle. Since we are sampling, uniformly,

$$p(\omega) = c \tag{47}$$

$$\int_{H^2} p(\omega) d\omega = 1 \Rightarrow c \int_{H^2} d\omega = 1 \Rightarrow c = \frac{1}{2\pi} \tag{48}$$

$$p(\omega) = \frac{1}{2\pi} \tag{49}$$

Therefore, $p(\theta, \phi) = \frac{1}{2\pi} \sin\theta$.
Let's sample θ first. Compute the marginal density:

$$p(\theta) = \int_0^{2\pi} p(\theta,\phi) d\phi = \int_0^{2\pi} \frac{\sin\theta}{2\pi} d\phi = \sin\theta \tag{50}$$

Now compute the conditional density:

$$p(\phi|\theta) = \frac{p(\theta,\phi)}{p(\theta)} = \frac{1}{2\pi} \tag{51}$$

This should make sense, since the hemisphere is symmetric around ϕ, so ϕ should really be sampled uniformly.

Now, just use the 1-D inversion method and integrate.

$$P(\theta) = \int_0^{\theta} P(\theta^*) d\theta^* = \int_o^{\theta} \sin\theta^* d\theta^* = 1 - \cos\theta \tag{52}$$

$$P(\phi|\theta) = \int_0^{\phi} p(\phi^*|\theta) d\phi^* = \frac{\phi}{2\pi} \tag{53}$$

And invert:

$$P(\theta) = U_1 \tag{54}$$

$$1 - \cos\theta = U_1 \tag{55}$$

$$\theta = \cos^{-1}(1 - U_1) \tag{56}$$

$$\theta = \cos^{-1}(U_1) \tag{57}$$

We can make the change in Equation **??** because if U_1 is uniform, then so is $1 - U_1$.

$$P(\phi|\theta) = U_2 \tag{58}$$

$$\frac{\phi}{2\pi} = U_2 \tag{59}$$

$$\phi = 2\pi U_2 \tag{60}$$

6.2 Example: Sampling a Unit Disk

Suppose we want to sample a unit disk uniformly with respect to area. This means that

$$p(x, y) = c \tag{61}$$

$$\int_{disk} c\,dx\,dy = 1 \Rightarrow c = \frac{1}{\pi} \tag{62}$$

So $p(x, y) = \frac{1}{\pi}$. By the transformation method,

$$p(r, \theta) = r p(x, y) = \frac{r}{\pi} \tag{63}$$

Now we compute the marginal density:

$$p(r) = \int_0^{2\pi} p(r, \theta) d\theta \tag{64}$$

$$p(r) = \int_0^{2\pi} \frac{r}{\pi} d\theta \tag{65}$$

$$p(r) = \frac{r}{\pi}(2\pi) \tag{66}$$

$$p(r) = 2r \tag{67}$$

Then we compute the conditional density:

$$p(\theta|r) = \frac{p(r, \theta)}{p(r)} \tag{68}$$

$$p(\theta|r) = \frac{\frac{r}{\pi}}{2r} \tag{69}$$

$$p(\theta|r) = \frac{1}{2\pi} \tag{70}$$

Notice Equation **??** is true again because of symmetry.
Next we integrate

$$P(r) = r^2 \tag{71}$$

$$P(\theta|r) = \frac{\theta}{2\pi} \tag{72}$$

Finally we invert:

$$P(r) = U_1 \tag{73}$$

$$r^2 = U_1 \tag{74}$$

$$r = \sqrt{U_1} \tag{75}$$

$$P(\theta|r) = U_2 \tag{76}$$

$$\frac{\theta}{2\pi} = U_2 \tag{77}$$

$$\theta = 2\pi U_2 \tag{78}$$

Note that we could have done this in the other order. Sometimes one order will be easier than the other, so if one way gets complicated, try the other.

6.3 A More Complex Example: Sampling a Triangle

Let's assume an isosceles right triangle (We are computing barycentric coordinates so this does not really matter) as shown in figure **??**.

$$p(u, v) = c \tag{79}$$

$$\int_{tri} c\,dx\,dy = 1 \tag{80}$$

$$c \int_{tri} dx\,dy = 1 \tag{81}$$

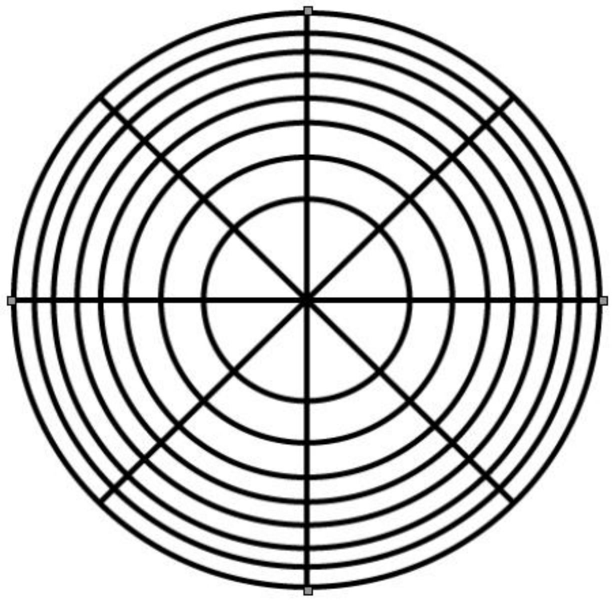

Figure 10: Sampling a circle uniformly

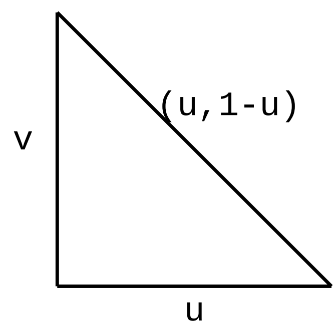

Figure 11: Sampling a triangle

$$c = \frac{1}{Area_{tri}} = 2 \tag{82}$$

So $p(u, v) = 2$. Now we compute the marginal density.

$$P(u) = \int_0^{1-u} P(u, v) dv \tag{83}$$

$$P(u) = 2 \int_0^{1-u} dv \tag{84}$$

$$P(u) = 2(1-u) \tag{85}$$

And the conditional density:

$$P(v|u) = \frac{(9u, v)}{P(u)} \tag{86}$$

$$P(v|u) = \frac{2}{2(1-u)} \tag{87}$$

$$P(v|u) = \frac{1}{1-u} \tag{88}$$

Now we integrate:

$$P(u) = \int_0^u 2(1-u^*)du^* = 2u - u^2 \tag{89}$$

$$P(v) = \int_0^v \frac{1}{1-u} dv = \frac{v}{1-u} \tag{90}$$

Finally we invert:

$$P(u) = U_1 \tag{91}$$

$$2u - u^2 = U_1 \tag{92}$$

$$u^2 - 2u + U_1 = 0 \tag{93}$$

$$u = \frac{2 \pm \sqrt{4 - 4U_1}}{2} \tag{94}$$

$$u = \frac{2 \pm 2\sqrt{1 - U_1}}{2} \tag{95}$$

$$u = 1 \pm \sqrt{1 - U_1} \tag{96}$$

$$u = 1 \pm \sqrt{U_1} \tag{97}$$

$$u = 1 - \sqrt{U_1} \tag{98}$$

We choose the negative root because $u \in [0, 1]$.

$$P(v) = U_2 \tag{99}$$

$$\frac{v}{1-u} = U_2 \tag{100}$$

$$v = U_2(1-u) \tag{101}$$

$$v = U_2(1-(1-\sqrt{U_1})) \tag{102}$$

$$v = U_2\sqrt{U_1} \tag{103}$$

Note that u and v are *not* independent!

6.4 Example: Cosine-weighted hemisphere

This comes up all the time when sampling the reflectance equation:

$$L_o(\omega_o) = \int_{H^2} f_r(\omega \to \omega_o) L_i(\omega) \cos\theta d\omega \tag{104}$$

Where

$L_o(\omega_o)$	Radiance reflected in direction ω_o
$f_r(\omega \to \omega_o)$	BRDF
$L_i(\omega)$	Incoming Radiance
θ	Angle between ω and N

If f_r is a constant, we can consider it to be like the OpenGL diffuse coefficient:

$$L_o(\omega_o) = \int_{H^2} K_D L_i(\omega) \cos\theta d\omega \tag{105}$$

We don't know anything about the incident illumination L_i, so we will try to sample according to a cosine distribution:

$$p(\omega) \propto \cos\theta \tag{106}$$

If we convert to spherical coordinates, we get:

$$p(\theta, \phi) \propto \sin\theta\cos\theta \tag{107}$$

Remember $\theta \in [0, \frac{\pi}{2}]$ and $\phi \in [0, 2\pi]$.
First, normalize:

$$\int_{H^2} cp(\theta,\phi) d\theta d\phi = 1 \tag{108}$$

$$\int_0^{2\pi}\int_0^{\frac{\pi}{2}} c(\cos\theta\sin\theta) d\theta d\phi = 1 \tag{109}$$

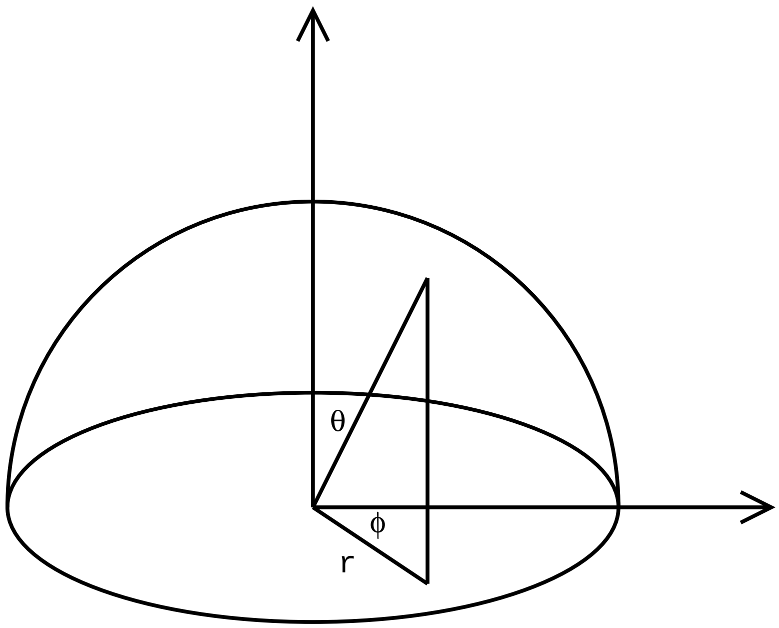

Figure 12: Polar Coordinates

$$c2\pi \int_0^1 u du = 1 \tag{110}$$

$$c2\pi \left(\frac{1}{2}\right) = 1 \tag{111}$$

$$c = \frac{1}{\pi} \tag{112}$$

Or

$$p(\theta, \phi) = \frac{1}{\pi} \cos\theta \sin\theta \tag{113}$$

We could marginalize and continue as before, but here's a trick: Choose a random point on the unit disk (uniformly with respect to area, as we just did), and project it upward onto the hemisphere!

Let (r, ϕ) be the polar coordinates of the point in the plane. (Note we are using ϕ instead of the usual θ). The joint density $p(r, \phi) = \frac{r}{\pi}$ represents the point sampled in the disk.

Now we map the hemisphere $\rightarrow$ vertical projection $\rightarrow \sin\theta = r$. This is easily seen from the diagram in figure **??**.

To complete the transformation $(r, \theta) \rightarrow (\sin\theta, \phi)$, we need the determinant of the Jacobian:

$$|J_T| = \left| \begin{array}{cc} \frac{\delta_r}{\delta_\theta} & \frac{\delta_r}{\delta_\theta} \\ \frac{\delta_\phi}{\delta_\theta} & \frac{\delta_\phi}{\delta_\theta} \end{array} \right| = \left| \begin{array}{cc} \cos\theta & 0 \\ 0 & 1 \end{array} \right| = \cos\theta \tag{114}$$

$$p(r,\theta) = \frac{1}{|J_T|} p(\theta,\phi) \Rightarrow p(\theta,\phi) = |J_T| p(r,\theta) = \cos\theta \frac{r}{\pi} = \frac{1}{\pi} \cos\theta \sin\theta \tag{115}$$

But this is what we wanted!

$$X = \sqrt{U_1} \cos(2\pi U_2) \tag{116}$$

$$Y = \sqrt{U_1} \sin(2\pi U_2) \tag{117}$$

$$Z = \sqrt{1 - (X^2 + Y^2)} = \sqrt{1 - (U_1(\cos^2 + \sin^2))} = \sqrt{1 - U_1} \tag{118}$$

Note this also works for other methods of sampling a unit disk (rejection, Shirley's mapping, etc).