Homework No. 2

Homeworks must be handed in by 5pm on the day they are due. You do not need to type it up but you will be penalized if we can't read it.

In this homework you're going to study Monte Carlo integration. In rendering, much of what we do is really integration. For instance, when you take a digital photograph, each pixel of the CCD sensor accumulates energy from incoming photons that arrive at that pixel which can be regarded as an integral of the incoming radiance. In our renderer, we'll have a radiance function that we want to integrate ... but evaluating this function will be quite expensive. That makes traditional Riemann integration out of the question.

For now, we're going to imagine that our "radiance function" is just a function defined on the unit interval



So the mean of the function is (A + B + C)/3. Problem 1 considers a way to estimate this value.

Problem 1

You need to estimate the mean value of f but you are only allowed to ask Spike one question ¹, namely, what is the value of f(x) for one specific x. So you do the following

Flip a 3-sided coin which lands on side i, where i is likely to be 0, 1, or 2. Ask Spike what is f(i/3).

Use that value as the estimate of the mean.

Because the algorithm uses random numbers, the estimate can be regarded as a random variable.

(a) What's the expected value, μ , of the random variable?

¹Note: please don't actually email Spike with these types of questions. We mean that you are only allowed to evaluate f at a single point. While this doesn't make a lot of sense in this example, if f were a complicated function and evaluating f(x) was computationally expensive, then the restriction becomes more applicable.

- (b) What's the variance? Express your answer in terms of A, B, C and μ .
- (c) Under what conditions on A, B and C is the variance 0? Under what conditions is it large compared to μ ?

Problem 2

Suppose we could generate random numbers uniformly in the unit interval. We could interpret problem 1 as saying "if the random number is between 0 and 1/3, pick A; if between 1/3 and 2/3, pick B; if between 2/3 and 1, pick C. Use the value as an estimate of the mean of A, B and C." We consider this as a 'single sample estimate of the mean.' We could, of course, take many samples and average them, and get a far better estimate. Later, we'll see quantitatively how much better.

Although it's very easy to generate points uniformly on an interval, it's not so easy to do so on some of the other domains of integration that we'll be looking at. So what happens if we have a random-number generator that picks points in the unit interval unevenly: 70% fall in the first third, 20% in the second third and 10% in the final third.



Could we still estimate the mean of f (i.e. (A+B+C)/3) given a single sample $x \in [0,1)$ generated from the above distribution and the corresponding value of f(x)?

- (a) Use the information above to derive an estimator ² for (A + B + C)/3
- (b) Compute the variance for your estimate.
- (c) When (qualitatively) would you expect this variance to be higher than the variance from problem 1? When would you expect it to be lower?

Problem 3

Same as for problem 2 but now, instead of estimating the mean of the function, we would like to estimate the value of the integral. Up until now, the function has been defined over the unit

²An estimator for a quantity *B* is a random variable Y, whose expected value should be close to (ideally equal to) *B* and whose variance should ideally be small. If E[Y] = B, then we call Y an unbiased estimator of *B*. The algorithm given in problem 1 is an example of an unbiased estimator with high variance.

Due: February 15th, 2010

interval so the mean and the integral happened to be the same- now let's see what happens if we expand things. We'll consider a function g, defined on the interval [0, 2] graphed below. and try to estimate its integral over [0,2] rather than [0,1]. The probability of drawing A, B, or C is the same as it was in problem 2.



- (a) How would you estimate the mean of g?
- (b) How would you estimate the integral of g?

Problem 4

Now imagine changing the number of values (there were 3 - A, B and C) to a larger number... perhaps infinitely many.



Spike would again like to know the average value of h.

- (a) Compute the average value of $h(x) = 1 + (\frac{x}{2})^2$ on the interval [0,4] using calculus.
- (b) Now, ESTIMATE the value by picking a random number x uniformly in [0,4] and evaluating h(x). Write a small program to do this computation 100 times and confirm that the mean of all the trials matches (closely) the value you got in part (a).(Don't need to hand the program in).
- (c) Now consider instead the function $h(x) = \left(\frac{x}{4}\right)^{10}$, whose average value on the interval is 1/11. Repeat part (b); how close is the mean to the actual average value. Try to explain why.

Part(c) shows that this *Monte Carlo* technique doesn't always produce a very accurate answer. One strategy for improving the technique is to draw samples from a non-uniform distribution. Intuitively, we'd like to favor points in places where the function is large and choose few points where the function is small. The goal of *importance sampling* is to reduce variance while approximating the integral of some function f over a domain D by drawing random samples across D according to some probability density function p which is proportional to f. By drawing more samples from the regions where f is large (concentrating our sampling effort on the "important" regions of the domain), and similarly drawing fewer samples from the regions where f is small (focusing less effort in the "unimportant" regions of the domain), the variance of our estimate overall is reduced, as long as we compensate for our uneven sampling rate.

Think of importance sampling this way: if we have only one shot at sampling f (only one sample because of very limited "resources"), we would like to concentrate that one sample in the region of the domain that will contribute the most to the value of the integral we are ultimately trying to approximate. By biasing our sampling technique towards "important", "large" regions of the domain with respect to f, we get more bang for our buck and correspondingly end up with a much lower variance estimator. If we were to instead naively sample uniformly across the domain, D, there's a good chance that our one, precious sample would be wasted by sampling a location, x that is unimportant, where f(x) is relatively small. We must compensate, however, for the fact that our estimate favors choosing certain values over others by dividing by something to keep our estimate unbiased.

We'll be using importance sampling in this class, but for now, we abandon this line of discussion and instead consider further problems in integration and random-number generation. As a first step, let's try to understand how our random estimates can be improved by computing multiple values and averaging.

Problem 5

Suppose X is a random variable drawn from a distribution p, and it has mean μ . What we've done in the previous problems is to estimate μ with X, and found the variance of our estimator in several cases. Now suppose that X_1 and X_2 are independent random variables distributed according to p, and we average them.

- (a) What's the variance of this average $\frac{X_1+X_2}{2}$? (You may need to look up rules for the variance of X + Y and cX).
- (b) What if instead we take the average of *n* random variables? Let $Z = \frac{1}{n} \sum_{i=1}^{n} X_i$, where $X_1...X_n$ are IID ³ random variables distributed according to *p*. What is the variance of *Z*?
- (c) Eplain what this means quitatively in terms of how accurate our estimator Z becomes as we increase the number of samples, n. (Note that standard deviation as opposed to variance is generally preferred for these types of questions because standard deviation has the same units as the mean which makes it easier to see the underlying relationship.)

³IID stands for Independent and Identically Distributed

Problem 6

Suppose you've got a random number generator that can generate random numbers uniformly in the interval [0, 1]. That means that if you draw a random number, X, then

$$Pr[a \le X \le b] = b - a$$

for $0 \le a < b, \le 1$. Alternatively, one can say that X is a random variable whose density is the contant function 1.

Now suppose that $f:[0,1] \to \mathcal{R}$ is a continuous, differentiable, increasing function (like $f(x) = x^2$). Suppose we define Y = f(X). Then Y will in general not be uniformly distributed on [f(a), f(b)], unless f happens to be an affine function (i.e., f(x) = px + q). To keep your notation simple, denote the *inverse* function of f by g, so that f(g(t)) = t and g(f(t)) = t, whenever these things make sense. And let's use c to denote f(a) and d = f(b), even though these might be infinite (i.e., f(x) could be something like 1/(1-x)).

As noted, g is not uniformly distributed in general.

- (a) Draw a graph showing a, b, c, d, and the graph of f for some continuous and increasing (but NOT affine) function f, so that you can refer to it for the rest of this exercise. Next to it, similarly labelled, draw the graph of g. Describe how these are related.
- (b) To compute the density for Y at a point p, we're going to first compute

$$Pr[p \le Y \le q],$$

and then let q = p + h, giving a formula for $Pr[p \le Y \le p + h]$. We'll divide this by h and let h go to zero, and that will give us the density. So: for a pair of numbers p < q in the range [c, d], express $Pr[p \le Y \le q]$ using f. Hint: Notice that because f is increasing, and hence 1-to-1, and the same is true of g, we know that $p \le f(X) \le q$ if and only if $g(p) \le g(f(X)) \le g(q)$. Work from there.

(c) Now use your formula from part b to write out an expression for

$$\frac{\Pr[p \le Y \le p+h]}{h}$$

and simplify this by approximating g(p+h) with Taylor's theorem at p.

- (d) Take the limit as h goes to zero to get an expression for the density of Y in terms of g or f.
- (e) Apply your formula to the special case of $f(x) = x^2$ to get an expression for the density for Y = f(X), where X is uniform on [0, 1].
- (f) Now suppose that you want to create a random variable Y on the interval $(\infty, 0]$, whose density at p is e^p , (This density clearly integrates to one on the negative reals). To do so, you need to choose a function f to apply to a uniform random variable on [0, 1]. Find the right function.

Problem 7

Given a uniform random number generator on the unit interval, how might we generate samples from an arbitrary distribution? This is an especially practical problem since most built-in computer random number generators only generate random numbers uniformly on the unit interval.

Specifically, suppose we want to generate points between 0 and 1 where the probability density of selecting x is, for example, $f_x(x) = 3x^2$ (which integrates to 1). There are a number of ways to do this; we'd like to use the following algorithm:

Draw X from a uniform distribution on [0, 1].

Draw another value s from a uniform distribution on [0,3] (3 is the largest value taken by f_x). If $s < f_x(x)$, keep x, otherwise throw it out and try again.

This is a particular instance of a strategy called "rejection sampling" (see Wikipedia) due to Neumann.

- (a) Given a uniform random number generator on [0,1], use rejection sampling to generate samples according to the probability density function $h_x(x) = \frac{3}{4}(1+x^2)$ What fraction of the samples do you expect to reject, on average, over a long run of trials? Implement; test that the distribution produced really looks like $h_x(x)$ (use hist in matlab) and verify your conjecture about the rejection rate.
- (b) What are some of the pros and cons of rejection sampling?

Problem 8

Now let's apply this to rendering: suppose that we know the radiance arriving at point P on some surface as a function of latitude and longitude on a hemisphere centered at P, one whose north-sourth axis is aligned with the normal n to the surface at P.



Let's call this incident randiance $L(\theta, \phi)$. We'd like to integrate L over the unit hemisphere to find the total incoming radiance, but since we tend to care about radiance per unit area, we need to account for the angle at which it arrives: we need to integrate $K(\theta, \phi) = L(\theta, \phi) \cdot \cos(\phi)$ where $\phi = 0$ is the north pole and $\phi = \pi/2$ is the equator (note that in class Spike parameterized $\phi = 0$ at the equator and $\phi = \pi/2$ at the north pole). For the remainder of this problem, let $L(\theta, \phi) = \frac{\cos(\phi)}{\sin(\phi)}$ and therefore $K(\theta, \phi) = \frac{\cos^2(\phi)}{\sin(\phi)}$.

Insight: Calculus is hard; instead, we'd like to *estimate* the integral of K over the hemisphere by generating random samples on the hemisphere and averaging K evaluated at those points. (Please handin Matlab code for all parts except part (a).

- (a) If we had a way to generate points uniformly on the hemisphere, one method for integrating K would be "compute K at each of the points and average the results." Imagine what would happen in this computation if K were the constant function whose value at every point was 1; what's the actual integral of K over the unit hemisphere? What's the value that our estimate would produce (independent of the number of points used)? What constant did we forget in our estimate and what does this mean?
- (b) A first attempt at generating random points uniformly on the hemisphere is to pick θ uniformly on $[0, 2\pi]$ and ϕ uniformly on $[0, \pi/2]$. Use Matlab to do this, and plot ⁴ 5000 such points. (To go from (θ, ϕ) to (x, y, z) use $x = \cos(\theta) \sin(\phi)$; $y = \sin(\theta) \sin(\phi)$; $z = \cos(\phi)$). Do the resulting points appear to be uniformly distributed?
- (c) As a cute hack from class, we can generate uniform samples using a unit cylinder. Generate 10,000 such samples and plot them by projecting them onto the unit sphere to see that they really *are* uniform, and then estimate the integral of K over the hemisphere using Monte Carlo integration with these samples (making sure to include the constant multiplier from part (a)). Record this value, as we'll be using it later to compare the values produced by other approaches.

The cute hack is nice, but let's imagine that we didn't have it from now on. How would we approache the problem? In part (d) we'll reject our way to make our samples uniform, and then parts (e) and (f) will explore how to estimate the integral more generically by compensating for the non-uniformity of the samples without actually changing the sample points. We'll compare these to the estimate we got with the cylindrically-generated samples in part (c), just to be sure everything worked out as expected.

(d) Now, use the rejection sampling technique from problem 7 to make your samples uniform. Once again, implement your approach in Matlab and compare your results to the other methods.

Part (d) can be viewed as evaluating the estimator $Y_1 = 2\pi \frac{1}{n} \sum_{i=1}^n K(\theta, \phi)$ where the joint distribution of θ and ϕ is uniform over the hemisphere. You should check for yourself to ensure that $E(Y_1)$ is what you'd expect. An equivalent way of writing this estimator is $Y_1 = \frac{1}{n} \sum_{i=0}^n \frac{K(\theta, \phi)}{\frac{1}{2\pi}}$. Why would we want to write it this way? Because it's the more general form of what's going on here; $\frac{1}{2\pi}$ is the joint probability density of θ and ϕ over the hemisphere, so what we've actually been evaluating is a special-case of: $Y_2 = \frac{1}{n} \sum_{i=0}^n \frac{K(\theta, \phi)}{h(\theta, \phi)}$ where $h(\theta, \phi)$ is the joint probability density with repect to θ and ϕ . Therefore, instead of adjusting your distribution to be uniform as we did in parts (d) and (e), you could alternatively compensate for the non-uniformity of the distribution using a standard Monte Carlo estimator similar to Y_2 (this is analogous to the approach we took in problem 2). Using the naive / intuitive mapping from part (b), along with the density of points it generates on the hemisphere, how can we estimate the integral of K without modifying our sample points at all? Part (e) aims to figure out the density $h(\theta, \phi)$, and part (f) aims to evaluate the resulting estimator.

 $^{^4}$ doc plot3

Due: February 15th, 2010

- (e) The area of the part of the sphere between the planes z = b and z = c depends only on b c(where $-1 \le b \le c \le 1$). Thus the set of points on the sphere between z = 0.3 and z = 0.4has the same area as the set of points between z = 0.1 and z = 0.2. (To see this is true, the surface area for a hemisphere with radius r is $A = r^2 \int_0^{2\pi} \int_0^{\pi/2} \sin(\phi) d\phi d\theta$. For slices of the hemisphere, change the limits of integration on the inner integral). Use this fact to determine the density h of your sample points for the approach taken in part (b) as a function of z; use Matlab's histogram function (doc hist) with bins of size 0.05 to plot this density as a function of x. What function of z describes this density? ⁵
- (f) Compensate for the non-uniformity of your samples by constructing a Monte Carlo estimator similar to Y_2 to compute the integral of incoming radiance, and numerically estimate the value of your estimator. Also compute your estimator's expected value using calculus to verify the correctness of your answer.
- (g) Compare your answere to parts (c) (d) and (f)- they should all be approximately the same. What are some of the pros and cons of the approaches from parts (c), (d), and (f)? When might one method be easier to implement or more efficient than the others, etc.?

⁵Hint: for the distribution generated in part (e) the density at a point (θ, ϕ) is $\frac{c}{\sin \phi}$ where $c = \frac{1}{\pi^2}$