PROBABILITY AND MONTE CARLO ESTIMATORS 6

CS184: COMPUTER GRAPHICS AND IMAGING

Feb 28 - Feb 29, 2024

1 Quick Terminology

Expectation: probability-weighted average of all possible values. In the discrete case, given by

$$E[X] = \sum_i x_i p_i$$

In the continuous case, given by

$$E[X] = \int xp(x)dx$$

.

Variance: the expected value of the squared deviation from the mean, or how spread apart values are from the mean. Given by

$$\begin{aligned} Var(x) &= E[(X - E[X])^2] \quad (1) \\ &= E[X^2] - E[X]^2 \quad (2) \end{aligned}$$

Cumulative Distribution Function (CDF): probability that a sample from distribution X will take a value less than or equal to x.

Lagrange Multipliers: a method to find the maxima or minima of a function $f(x)$ with constraints $g(x) = 0$. We create a function

$$L(x, \lambda) = f(x) + \lambda g(x)$$

and look for critical points where the gradient of L is 0. The critical points of L are the maxima/minima of f.

2 Inversion Method

Recall the inversion method from class. Given a uniform random variable U in the interval $[0, 1]$, we can generate a random variable from any other one dimensional distribution if we have access to the inverse of its cumulative distribution function, $F^{-1}(x)$. We simply have to return $X = F^{-1}(U)$. This is how we choose sample points when running a ray tracing algorithm.

1. What function of U will return a sample from the exponential distribution (with parameter λ)? This distribution has density

$$p_\lambda(x) = \lambda e^{-\lambda x}$$

and is defined for $x \geq 0$.

Solution: First, we need to calculate the CDF:

$$F_\lambda(x) = \int_0^x p_\lambda(x)dx = \int_0^x \lambda e^{-\lambda x}dx = -e^{-\lambda x}\Big|_0^x = 1 - e^{-\lambda x}$$

Its inverse is

$$F_\lambda^{-1}(x) = -\frac{\log(1-x)}{\lambda}$$

Thus we can return

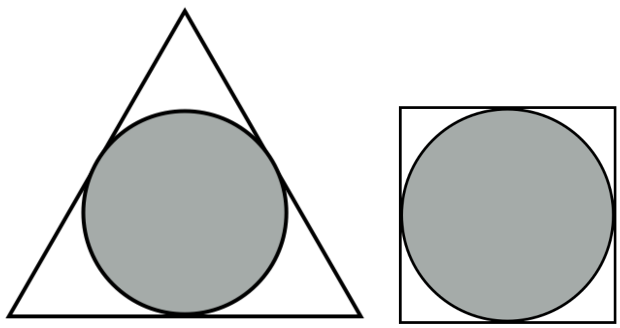
$$-\frac{\log(1-U)}{\lambda}$$

to sample from the exponential distribution.

3 Rejection Sampling

Recall that rejection sampling is one way of using the Monte Carlo method to sample from a probability distribution. We repeatedly sample values from a proposed distribution, then accept or reject that sample based on whether or not it falls within a probability density function that we know how to sample from. If the sample falls outside of the PDF, then we reject that sample. The remaining samples should be uniformly distributed within our target probability function.

The figure below shows two methods for estimating π using rejection sampling. Method (a) generates random points uniformly in an equilateral triangle. Method (b) generates random points uniformly in a square. Both methods estimate π using the ratio of the number of points that lie within the shape's tangent circle to the total number of points sampled.



(a) Triangle (b) Square

1. In method (a), suppose there are k points out of all n points sampled that lie within the triangle's tangent circle. Then, a formula for the estimated value of π is:

Solution: We can relate the ratio between k and n to the ratio between the area of the circle and the area of the triangle. Note that the area of an equilateral triangle with an inscribed circle of radius r is $3\sqrt{3}r^2$. Set up the ratio and then solve for π.

$$\frac{k}{n} = \frac{\pi r^2}{3\sqrt{3}r^2}$$

$$\pi = \frac{3\sqrt{3}k}{n}$$

2. In method (b), suppose there are k points out of all n points sampled that lie within the square's tangent circle. Then, a formula for the estimated value of π is:

Solution: As above, we relate the ratio between k and n to the ratio between the area of the circle and the area of the square. The square with an inscribed circle of radius r has sides of length $2r$.

$$\frac{k}{n} = \frac{\pi r^2}{4r^2}$$
$$\pi = \frac{4k}{n}$$

3. Which method ((a) or (b)) has lower variance with the same number of samples?

Solution: (b) has lower variance, because a larger percentage of the points will land in the shaded area.

4. What are some downsides to using uniform random sampling?

Solution: It may take a lot of samples to get accurate representation of desired distribution, especially if the sample points lie on one side of the distribution. To fix this problem, we want to sample more intelligently – see importance sampling.

Optional: there's also a notion of discrepancy – from Wikipedia, the discrepancy of a sequence is low if the proportion of points in the sequence falling into an arbitrary set B is close to proportional to the measure of B. With that in mind, in many cases people use low-discrepancy sequences to sample. When these sequences are used to numerically approximate an integral, the method is called quasi-Monte Carlo integration.

5. What are some downsides to using rejection sampling for high-dimensional spaces?

Solution: We run into the problem commonly known as the "curse of dimensionality," which means that many more samples will be rejected than will be accepted. This means that the efficiency of rejection sampling is dramatically reduced in higher-dimensional spaces.

4 Unbiased Monte Carlo Estimator

The Monte Carlo method can also be used to estimate an integral. When this method is used properly, the expectation of an **unbiased** Monte Carlo estimator is equal to the true value of the integral.

1. You have two random variables X and Y, which are drawn uniformly from $[-2, 2]$. What is an **unbiased** Monto Carlo estimator for the given integral?

$$F = \int_{-2}^{2}\int_{-2}^{2} f(x, y)dxdy$$

(Hint: Your answer should include a summation over N samples.)

Solution: The Monte Carlo estimator is:

$$\begin{aligned}\langle F_N\rangle \frac{1}{N}\sum_{i=1}^{N}\frac{f(X_i, Y_i)}{p(X_i, Y_i)} &\text{, where } p(X_i, Y_i) = p(X_i)p(Y_i) = (\frac{1}{4})(\frac{1}{4}) = \frac{1}{16} \\ = \frac{1}{N}\sum_{i=1}^{N}\frac{f(X_i, Y_i)}{\frac{1}{16}} & \\ = \frac{16}{N}\sum_{i=1}^{N} f(X_i, Y_i) &\end{aligned}$$

The notation $\langle F_N\rangle$ here represents an approximation of F using N samples.