

Homework No.5, 550.695, Due November 22, 2011.

1. This problem discusses the inverse sampling method for several examples.

(a) The *stretched-exponential PDF* has the form

$$p_{\beta,\gamma,\lambda}(x) = \frac{\beta\lambda^{(\gamma+1)/\beta}}{\Gamma\left(\frac{\gamma+1}{\beta}\right)} x^\gamma e^{-\lambda x^\beta}, \quad x \geq 0$$

and $= 0$ for $x < 0$. Find the corresponding CDF in terms of the incomplete Gamma function. For the case $\beta = 1/\sqrt{5}$, $\gamma = \pi$, $\lambda = 7$, sample this distribution by the inverse sampling method using `gammaincinv` in `MATLAB`. Produce an empirical PDF with $x \in [0, 10]$, using $dx = 0.1$ and $N = 10^7$ and compare with the exact PDF. Plot the two PDF's in linear-linear coordinates and also the logarithm of the two PDF's in log-log coordinates. In the latter you should see at large x a line with slope $-\beta$.

(b) A special case for $\gamma = 0$, $\beta = 1$ is the *exponential random variable* $\text{Exp}(\lambda)$. Show that

$$\tilde{X} = -\ln(\tilde{U})/\lambda$$

is $\text{Exp}(\lambda)$ if \tilde{U} is a uniform random variable.

(c) The *Poisson random variable* $\tilde{N} \sim \text{Poisson}(\lambda)$ with intensity parameter λ takes on non-negative integer values with probabilities

$$P(\tilde{N} = k) = e^{-\lambda} \frac{\lambda^k}{k!},$$

for $k = 0, 1, 2, \dots$. Use the results of Homework #2, Problem 2 to generate i.i.d. $\text{Poisson}(\lambda)$ random variables in terms of an i.i.d. sequence of $\text{Exp}(\lambda)$ random variables \tilde{X}_k , $k = 1, 2, 3, \dots$ by setting

$$\tilde{N} = \max\{k : \tilde{X}_1 + \tilde{X}_2 + \dots + \tilde{X}_k \leq 1\}$$

and $\tilde{N} = 0$ if $\tilde{X}_1 > 1$. Produce a (discrete) PDF at $k = 0, 1, \dots, 15$ for a Poisson r.v. with $\lambda = 4$ using $N = 10^6$ samples and compare with the exact PDF.

2. This problem discusses the Metropolis and Metropolis-Hastings methods to sample random vectors.

(a) Consider the angle distribution on $-\pi < \theta < \pi$

$$p(\theta) = \frac{1}{2\pi I_0(\beta)} \exp(\beta \cos \theta),$$

for real β and $I_0(\beta)$ the modified Bessel function. This distribution appears in various problems. For example, it is the stationary distribution of an overdamped pendulum driven by a Brownian motion. It also appears in the theory of “stochastic resonance” (e.g. see L. Gammaitoni et al., Rev. Mod. Phys., Volume 70, pp. 223 - 287 (1998).) Sample this distribution for $\beta = 1$ using the Metropolis method. Produce an empirical PDF using $d\theta = \pi/50$ and $N = 10^7$ and compare with the exact PDF. What is the acceptance rate of the Metropolis algorithm?

(b) As an approximation to the doublewell PDF

$$p(x) = \frac{1}{Z} \exp\left(-\frac{U(x)}{\kappa}\right), \quad U(x) = -\frac{1}{2}x^2 + \frac{1}{4}x^4$$

consider the *Gaussian mixture*

$$q(x) = \frac{1}{2} \left[\frac{1}{\sqrt{\pi\kappa}} \exp\left(-\frac{(x+1)^2}{\kappa}\right) + \frac{1}{\sqrt{\pi\kappa}} \exp\left(-\frac{(x-1)^2}{\kappa}\right) \right]$$

Show by Taylor expansions around $x = \pm 1$ that $U(x) = -1/4 + (x \pm 1)^2 + O((x \pm 1)^3)$, so that the Gaussian terms match the doublewell PDF to quadratic order at the peaks. (In fact, it can be shown by the Laplace method of steepest descent that such a Gaussian mixture model becomes asymptotically exact for the limit $\kappa \rightarrow 0$.) Explain how to generate samples from the Gaussian mixture by choosing at random either the left or right Gaussian component and then sampling the corresponding normal vector. Use this result to sample from the doublewell PDF $p(x)$ for $\kappa = 1$ by the Metropolis-Hastings method with the Gaussian mixture $q(x)$ as a proposal distribution. Compare the exact PDF $p(x)$ and the empirical PDF $p^{(N)}(x)$ for $N = 10^8$, with $x \in [-4, 4]$ and bins of size $dx = 0.1$. Also compare this MCMC method with the Metropolis method discussed in class, in particular in terms of the acceptance rate of proposals for the two methods. Can you explain in intuitive terms the different acceptance rates?

3. The method of *importance sampling* is very useful when a Monte Carlo average

$$\langle f(\tilde{X}) \rangle_N = \frac{1}{N} \sum_{n=1}^N f(\tilde{X}_n)$$

to calculate the mean $\langle f(\tilde{X}) \rangle = \int dx f(x)p_X(x)$ is dominated by rare events. This makes random fluctuations large and convergence slow. The idea of importance sampling is to introduce a new distribution $p'_X(x) = w(x)p_X(x)$ where the weight $w(x) > 0$ satisfying $\int dx w(x)p_X(x) = 1$ is chosen to make very probable the rare events that contribute most to the average. Then,

$$\langle f(\tilde{X}) \rangle = \int dx \frac{f(x)}{w(x)} p'_X(x) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \frac{f(\tilde{X}'_n)}{w(\tilde{X}'_n)},$$

where \tilde{X}'_n are sampled from the distribution $p'_X(x)$. If f is positive, then the best choice of $w(x)$ is $f(x)/\langle f(\tilde{X}) \rangle$, but this requires knowing the very average $\langle f(\tilde{X}) \rangle$ that one is trying to calculate! One must try to find a $w(x)$ which is close to this optimal choice but which can be shown to satisfy the normalization condition $\int dx w(x)p_X(x) = 1$.

(a) Calculate $\langle \exp(3\tilde{X}^{21/20}) \rangle$ for $\tilde{X} \sim N(0, 1)$, where $x^{21/20}$ means here precisely $\text{sign}(x)|x|^{21/20}$. Use first direct averaging over i.i.d. $N(0, 1)$ samples and then improve by importance sampling with $w(x) = e^{3x-9/2}$. Show that w is normalized and that \tilde{X}' is an $N(3, 1)$ random variable. Compare the results of the direct and importance-sampling methods for $N = 10^6, 10^7, 10^8, 10^9$. Which method converges faster?

(b) Now calculate $\langle \exp(\frac{7}{16}|\tilde{X}|^{25/13}) \rangle$ for $\tilde{X} \sim N(0, 1)$ using first direct averaging over i.i.d. $N(0, 1)$ samples and then improve by importance sampling with a $w(x)$ of your choice. By increasing N in powers of 10, get the average converged to at least four significant figures using the importance sampling technique. State your choice of weight $w(x)$, the values of N considered and your result for the average at each N .