

Homework No.5, 550.695, Due April 2, 2009.

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

(a) The *Exponential*(b) random variable $\tilde{X} \sim \text{Exp}(b)$ for real $b > 0$ is a nonnegative random variable with the PDF

$$p_b(x) = be^{-bx}, \quad x \geq 0$$

and $p_b(x) = 0$ for $x < 0$. Show that

$$\tilde{X} = -\frac{1}{b} \ln(\tilde{U})$$

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

(b) The *Cauchy*(a, b) random variable $\tilde{X} \sim \text{Cauchy}(a, b)$ for real a, b with $b > 0$ has the PDF

$$p_{a,b}(x) = \frac{b}{\pi} \frac{1}{b^2 + (x - a)^2}.$$

Calculate the corresponding CDF. Use this result to write a code that generates an i.i.d. sequence of $\text{Cauchy}(a, b)$ random variables by the inverse sampling method. Use this code to produce an empirical PDF for $a = 0, b = 1$ with $N = 10^6$ samples and with bins of width $dx = 0.1$ for $x \in [-10, 10]$. Compare with the exact PDF.

(c) The *Poisson*(λ) 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 $k = 0, 1, 2, \dots$. Write a code to generate i.i.d. $\text{Poisson}(\lambda)$ random variables by the inverse sampling method. Produce a (discrete) PDF at $k = 0, 1, \dots, 15$ for a Poisson r.v. with $\lambda = 5$ using $N = 10^6$ samples and compare with the exact PDF.

Hint: Use the formula $F(x) = \sum_{k=0}^{[x]} e^{-\lambda} \frac{\lambda^k}{k!}$ for the CDF, where $[x]$ is the greatest integer less than or equal to x .

2. This problem discusses the bivariate normal distribution and its sampling.

(a) For any pair of random variables \tilde{X}, \tilde{Y} one defines standard deviations σ_X^2, σ_Y^2 by

$$\sigma_X^2 = \text{var}(\tilde{X}) = E(\tilde{X}^2) - (E\tilde{X})^2 = E\left[(\tilde{X} - E\tilde{X})^2\right]$$

and similarly for σ_Y^2 , and the correlation coefficient $\rho_{X,Y}$ by

$$\rho_{X,Y} = \frac{\text{cov}(\tilde{X}, \tilde{Y})}{\sigma_X \sigma_Y},$$

with $\text{cov}(\tilde{X}, \tilde{Y}) = E(\tilde{X}\tilde{Y}) - E(\tilde{X})E(\tilde{Y})$ the cross-covariance. Use the Cauchy-Schwartz inequality to show that $|\rho_{X,Y}| \leq 1$.

(b) For the random vector $\tilde{\mathbf{V}} = [\tilde{X} \ \tilde{Y}]^\top$ show that the covariance matrix

$$\mathbf{C} = E(\tilde{\mathbf{V}}\tilde{\mathbf{V}}^\top) - E(\tilde{\mathbf{V}})E(\tilde{\mathbf{V}})^\top = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$$

with $\rho = \rho_{X,Y}$. Then show that the PDF of a normal random vector $\tilde{\mathbf{V}}$ with zero mean and covariance \mathbf{C} , or $p_{\mathbf{V}}(\mathbf{v}) = \frac{1}{2\pi\sqrt{\text{Det}\mathbf{C}}} \exp\left(-\frac{1}{2}\mathbf{v}^\top\mathbf{C}^{-1}\mathbf{v}\right)$ with $\mathbf{v} = (x, y)$, is given by

$$p_{\mathbf{V}}(\mathbf{v}) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_X^2} - 2\rho\frac{xy}{\sigma_X\sigma_Y} + \frac{y^2}{\sigma_Y^2}\right)\right].$$

This is called the *bivariate normal density*.

(c) Find the lower-triangular Cholesky factor

$$\mathbf{L} = \begin{bmatrix} \ell_{XX} & 0 \\ \ell_{XY} & \ell_{YY} \end{bmatrix}$$

satisfying $\mathbf{C} = \mathbf{L}\mathbf{L}^\top$ explicitly in terms of σ_X, σ_Y , and ρ .

(d) Use the results in (c) to write a code to generate i.i.d. samples by representing the normal random vector as $\tilde{\mathbf{V}} = \mathbf{L}\tilde{\mathbf{N}}$ where $(\tilde{N}_X, \tilde{N}_Y)$ are independent standard normal r.v.'s. Use your code to simulate the bivariate normal variables \tilde{X}, \tilde{Y} with $\sigma_X = 1, \sigma_Y = 1/\sqrt{2}$ and $\rho = 1/\pi$ using $N = 10^8$ samples. Calculate the empirical joint PDF $p_{X,Y}^{(N)}(x, y)$ for $(x, y) \in [-4, 4]^2$ with $dx = dy = 0.1$ and compare with the exact PDF in (b). For example, plot both in MATLAB using `contour` or `surf`.

3. This problem discusses the Metropolis-Hastings method to sample random vectors.

(a) Recall that the *Gamma*(a, b) random variable $\tilde{X} \sim \Gamma(a, b)$ for real a, b with $b > 0$ is a nonnegative random variable with the PDF

$$p_{a,b}(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}, \quad x \geq 0$$

and $p_{a,b}(x) = 0$ for $x < 0$. Use the results in Problem 1(a) to write a code to sample the *Gamma*(a, b) distribution by the Metropolis-Hastings method, with $\text{Exp}(b)$ as the proposal distribution. Produce an empirical PDF for the $\Gamma(3, 1)$ r.v. with $x \in [0, 8]$, using $dx = 0.1$ and $N = 10^7$ and compare with the exact PDF. How do you expect the proposal acceptance rate in Metropolis-Hastings depends upon the parameter a ?

(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?