STAT 525 | Homework 6

Spring 2020 | Mar 12

Homework 6

Due on Thursday, March 26.

1. Design a Metropolis-Hastings algorithm to simulate samples from a distribution with the following density:

$$\pi(x) \propto e^{-x^6 - x}, \qquad -\infty < x < \infty.$$

Estimate the acceptance rate of your algorithm. Show the autocorrelation plot of your samples. Estimate the mean of the above density function and give the standard error of your estimate. Attach your code and results.

2. (Diaconis and Sturmfels, 1985) The following table shows data gathered to test the hypothesis of association between birthday and deathday. The table records the month of birth and death for 82 descendants of Queen Victoria. A widely stated claim is that birthday-deathday pairs are associated. The table is available in the homework web page.

Month of birth	Month of death												
	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec	Total
Jan	1	0	0	0	1	2	0	0	1	0	1	0	6
Feb	1	0	0	1	0	0	0	0	0	1	0	2	5
March	1	0	0	0	2	1	0	0	0	0	0	1	5
April	3	0	2	0	0	0	1	0	1	3	1	1	12
May	2	1	1	1	1	1	1	1	1	1	1	0	12
June	2	0	0	0	1	0	0	0	0	0	0	0	3
July	2	0	2	1	0	0	0	0	1	1	1	2	10
Aug	0	0	0	3	0	0	1	0	0	1	0	2	7
Sept	0	0	0	1	1	0	0	0	0	0	1	0	3
Oct	1	1	0	2	0	0	1	0	0	1	1	0	7
Nov	0	1	1	1	2	0	0	2	0	1	1	0	9
Dec	0	1	1	0	0	0	1	0	0	0	0	0	3
Total	13	4	7	10	8	4	5	3	4	9	7	8	82

Implement the exact test using the Metropolis-Hastings algorithm to test the null hypothesis that birthday and deathday are independent. One way to define p-value for exact test is

$$p\text{-value} = \sum_{T \in \Omega} 1_{\{p(T) \le p(T_0)\}} p(T),$$

where T_0 is the observed table above, and Ω is the set of 12×12 contingency tables with the same row sums and column sums as T_0. Here $p(T)$ is defined as

$$p(T) \propto \frac{1}{\prod_{i=1}^{12} \prod_{j=1}^{12} t_{ij}!}, \qquad T \in \Omega,$$

where t_{ij} is the (i, j)-th entry of the table. Estimate the p-value and give the standard error of your estimate by the Metropolis-Hastings algorithm given in class, which can generate samples approximately from the target distribution $p(T)$ on the tables in Ω:

At each step, pick two rows and two columns uniformly at random. Then adding or subtracting one in the four entries at the intersection of the two rows and two columns according to the following pattern:

$$\begin{array}{cc} + & - \\ - & + \end{array} \quad \text{or} \quad \begin{array}{cc} - & + \\ + & - \end{array} .$$

The two patterns are chosen with equal probability. Attached your code and results.