Computational Statistics 732A89 – Spring 2026
Computer Lab 4

Bayu Brahmantio, Xiaochen Liu, Frank Miller
Department of Computer and Information Science (IDA), Linköpings University

February 10, 2026

This computer laboratory is part of the examination for the Computational Statistics course. Create a group report (which is directly presentable, if you are a presenting group), on the solutions to the lab as a PDF file. Be concise and do not include unnecessary printouts and figures produced by the software and not required in the assignments. All R code should be included as an appendix to your report. A typical lab report should contain 2-4 pages of text plus some figures plus the appendix with codes.
In the report, refer to all consulted sources and disclose all collaborations. The report has to be written in English.
The report should be handed in via LISAM (or alternatively in case of problems by email) by **23:59 February 17, 2026** at the latest. Notice that there is a deadline for corrections 23:59 07 April 2026 and a final deadline of 23:59 28 April 2026 after which no submissions or corrections will be considered, and you will have to redo the missing labs next year.
The seminar for this lab will take place **February 25, 2026**.

Question 1: Computations with Metropolis–Hastings

Consider a random variable X with the following probability density function:

$$f(x) = \frac{1}{2.4 \cdot 10^6} x^4 e^{-x/10}, \quad x > 0.$$

a. Use the Metropolis–Hastings algorithm to generate 10000 samples $X_t, t = 1, \ldots, 10000$, from this distribution. Use a normal distribution with mean X_t and standard deviation 1 as proposal distribution when you generate X_{t+1}; take some starting point. Plot the chain you obtained with iterations on the horizontal axis. What does this plot say about the convergence of the chain? Is a burn–in period needed and what can be the size of this period? What is the acceptance rate? Plot a histogram of the sample.

b. Perform Part a by using the chi–square distribution with $\lfloor X_t + 1 \rfloor$ degrees of freedom as a proposal distribution, where $\lfloor x \rfloor$ is the floor function, meaning the integer part of x for positive x, i.e. $\lfloor 2.95 \rfloor = 2$.

c. Suggest another proposal distribution (can be a normal or chi–square distribution with other parameters or another distribution) with the potential to generate a good sample. Perform part a with this distribution.

d. Compare the results of Parts a, b, and c and make conclusions.

e. Estimate $E(X) = \int_0^\infty x f(x)\, dx$ and the median of X using the samples from Parts a, b, and c.

f. The distribution generated is in fact a gamma distribution. Search in the literature/internet and find out the exact value of the integral $E(X) = \int_0^\infty x f(x)\, dx$. Compare it with the one you obtained.

Question 2: Gibbs sampling

Let $X = (X_1, X_2)$ be a bivariate distribution with density $f(x_1, x_2) \propto \mathbf{1}\{x_1^2 + wx_1x_2 + x_2^2 < 1\}$ for some specific w with $|w| < 2$. X has a uniform distribution on some two-dimensional region. We consider here the case $w = 1.999$ (in Lecture 4, the case $w = 1.8$ was shown).

a. Draw the boundaries of the region where X has a uniform distribution. You can use the code provided on the course homepage and adjust it.

b. What is the conditional distribution of X_1 given X_2 and that of X_2 given X_1?

c. Write your own code for Gibbs sampling the distribution. Run it to generate $n = 1000$ random vectors and plot them into the picture from Part a. Determine $P(X_1 > 0)$ based on the sample and repeat this a few times (you need not to plot the repetitions). What should be the true result for this probability?

d. Discuss, why the Gibbs sampling for this situation seems to be less successful for $w = 1.999$ compared to the case $w = 1.8$ from the lecture.

e. We might transform the variable X and generate $U = (U_1, U_2) = (X_1 - X_2, X_1 + X_2)$ instead. In this case, the density of the transformed variable $U = (U_1, U_2)$ is again a uniform distribution on a transformed region (no proof necessary for this claim). Determine the boundaries of the transformed region where U has a uniform distribution on. You can use that the transformation corresponds to $X_1 = (U_2 + U_1)/2$ and $X_2 = (U_2 - U_1)/2$ and set this into the boundaries in terms of X_i. Plot the boundaries for (U_1, U_2). Generate $n = 1000$ random vectors with Gibbs sampling for U and plot them. Determine $P(X_1 > 0) = P((U_2 + U_1)/2 > 0)$. Compare the results with Part c.