



Contact during exam:
Finn Lindgren 91897822

Exam for TMA4300 Computationally intensive statistical methods
(Beregningskrevende statistiske metoder)
Friday June 10, 2011
Time: 09:00–13:00

ENGLISH

Permitted material:

- Calculator with empty memory.
- K. Rotman: *Matematisk formelsamling*.
- *Statistiske tabeller og formler*, Tapir.
- One yellow paper (A4 with stamp from the department of mathematical sciences) with your own formulas and notes.

All answers should be motivated.

You may answer in English or Norwegian.

Du kan besvare enten på engelsk eller norsk (begge målføre).

Problem 1

Let $f_X(x) = 2/(2+x)^2$, $x \geq 0$, be the probability density function of a random variable X . Assume that you have a procedure given for generating independent random variables U with uniform density on the interval $(0, 1)$.

- a) Describe an algorithm which generates random samples from the distribution for X , using the inverse cumulative probability function.
- b) Construct a rejection sampler for the density $f_Y(y) = \exp(-y)$, $y \geq 0$, which uses samples of X .
- c) What is the expected number of samples needed from X required to generate 1000 samples from Y ? On average, how many samples (in total) are required from U ?

Problem 2

Let X be a p -dimensional random variable with probability density function $f_X(x) > 0$ for all x .

- a) Gibbs sampling is one MCMC algorithm that can be used to generate samples from $f_X(x)$. Describe the Gibbs-sampling algorithm in general.
- b) Describe, in general terms, how you would verify that your Gibbs-sampler has (practically) converged.
- c) Your friend wants to estimate $E(X_1)$ as the mean of n samples of X_1 found from a Gibbs-sampler algorithm. Your friend tells you that (s)he would like to use an ordinary Student- t test to check if her/his estimate supports the claim that $E(X_1) = 0$. Explain why your friend cannot use the ordinary Student- t test for this purpose.
- d) Assume you have a Gibbs-sampler algorithm that sample from $f_X(x)$, and that you are interested in the probability

$$q = \text{Prob} \left(X_1 > \frac{X_2 + X_3}{2} \right).$$

Explain how you could estimate q using output from the Gibbs-sampler.

Problem 3 [This problem has double weight.]

Chose one of the following two topics,

1. cross-validation, or
2. bootstrapping.

Write an introduction to topic which includes at least one example of its usage.