NTNU
Norwegian University of
Science and Technology

Department of Mathematical Sciences

Examination paper for
TMA4300 Computer Intensive Statistical Methods

Academic contact during examination: Sara Martino

Phone: 9940 3330

Examination date: 05/06/2019

Examination time (from–to): 15.00-19.00

Permitted examination support material: C:

- Calculator HP30S, CITIZEN SR-270X or CITIZEN SR-270X College, Casio fx-82ES PLUS with empty memory.
- Statistiske tabeller og formler, Tapir.
- K. Rottmann: Matematisk formelsamling.
- One yellow, stamped A5 sheet with own handwritten formulas and notes.
- Dictionary in any language.

Other information:

- All 9 sub-problems in this exam count approximately the same.
- All answers must be justified!!!
- In your solution you can use English and/or Norwegian.

Language: English

Number of pages: 5

Number of pages enclosed: 0

Checked by:

Informasjon om trykking av eksamensoppgave	
Originalen er:	
1-sidig ☐	**2-sidig** ☐
sort/hvit ☐	**farger** ☐
skal ha flervalgskjema ☐	

Date Signature

Problem 1

Assume you are only able to sample from a uniform distribution Unif(0,1).

a) We want to draw samples from the following distribution:

$$g(x) = \frac{\lambda}{2} \exp(-\lambda|x|), \quad x \in \mathcal{R}$$

Describe how you can simulate from this distribution by one of the simulation methods we have discussed in Part 1 of this course. In particular, specify what method you choose to use, develop mathematical expressions necessary to implement the simulation method and write pseudo-code for generating one sample from the distribution.

In the following you can assume that in addition to the standard uniform distribution, you are also able to sample from the distribution considered in **a)**.

b) We want to generate samples from a standard normal distribution with density

$$h(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \quad x \in \mathcal{R}$$

Describe how you can simulate from this distribution by rejection sampling, using $g(x)$ specified in **a)** as proposal distribution. Develop mathematical expressions necessary to implement the rejection sampling algorithm in this case and write pseudo-code for generating one sample from the distribution.

c) Derive the overall acceptance rate of the rejection algorithm specified in **b)**. Which value of λ should you choose to optimize your algorithm? And what is the maximum obtainable acceptance rate?

Problem 2

In this problem we will consider a hierarchical Bayesian model for analysing data from an inhomogeneous Poisson process that we observe in discrete time.

Let Y_i, $i = 1, \ldots, n$, be the number of events occurring in the time interval $[i-1, i]$ and assume that

$$P(Y_i = y_i|\lambda, \theta, k) = \begin{cases} \frac{\lambda^{y_t}}{y_i!} e^{-\lambda} & \text{for } i \leq k, \\ \\ \frac{\theta^{y_t}}{y_i!} e^{-\theta} & \text{for } i > k, \end{cases}$$

where k is an integer value. In other words, the intesity of the process is λ before time k and θ after time k.

We are interested in estimating the two intensity parameters λ and θ and the change point of the process k. We assume that these parameters are apriori mutually independent and we assign the following prior distributions:

$$\theta \sim \text{Gamma}(0.5, 1), \;\; \lambda \sim \text{Gamma}(0.5, 1), \;\; k \sim \text{Unif}(1, 2, 3, \ldots, n)$$

Assume the Gamma density parameterisation

$$\text{Gamma}(\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\beta x}.$$

a) Derive the posterior density of the unknown parameters given the observed data $f(\theta, \lambda, k|y_1, \ldots, y_n)$. Derive also the full conditionals for the three parameters. You need to derive these densities only up to a proportionality constant. If possible specify what parametric family each full conditional belongs to, and their parameter values.

b) Use pseudo code to outline how you would generate samples from the posterior distribution using Markov chain Monte Carlo (MCMC). Specify in particular what your proposal distributions are and simplify as much as possible the expressions for the corresponding acceptance probabilities.

c) Figure 1 shows the trace plot of the first 2500 iterations of an MCMC algorithm for the above model where $n = 350$. What can you say about the convergence and the mixing properties by looking at these plots?

Explain how you can use the MCMC output to estimate the following quantities:

- The posterior mean of λ: $E[\lambda|y_1, \ldots, y_n]$
- The conditional probability $P(\lambda > 2.5|k = 100, y_1, \ldots, y_n)$

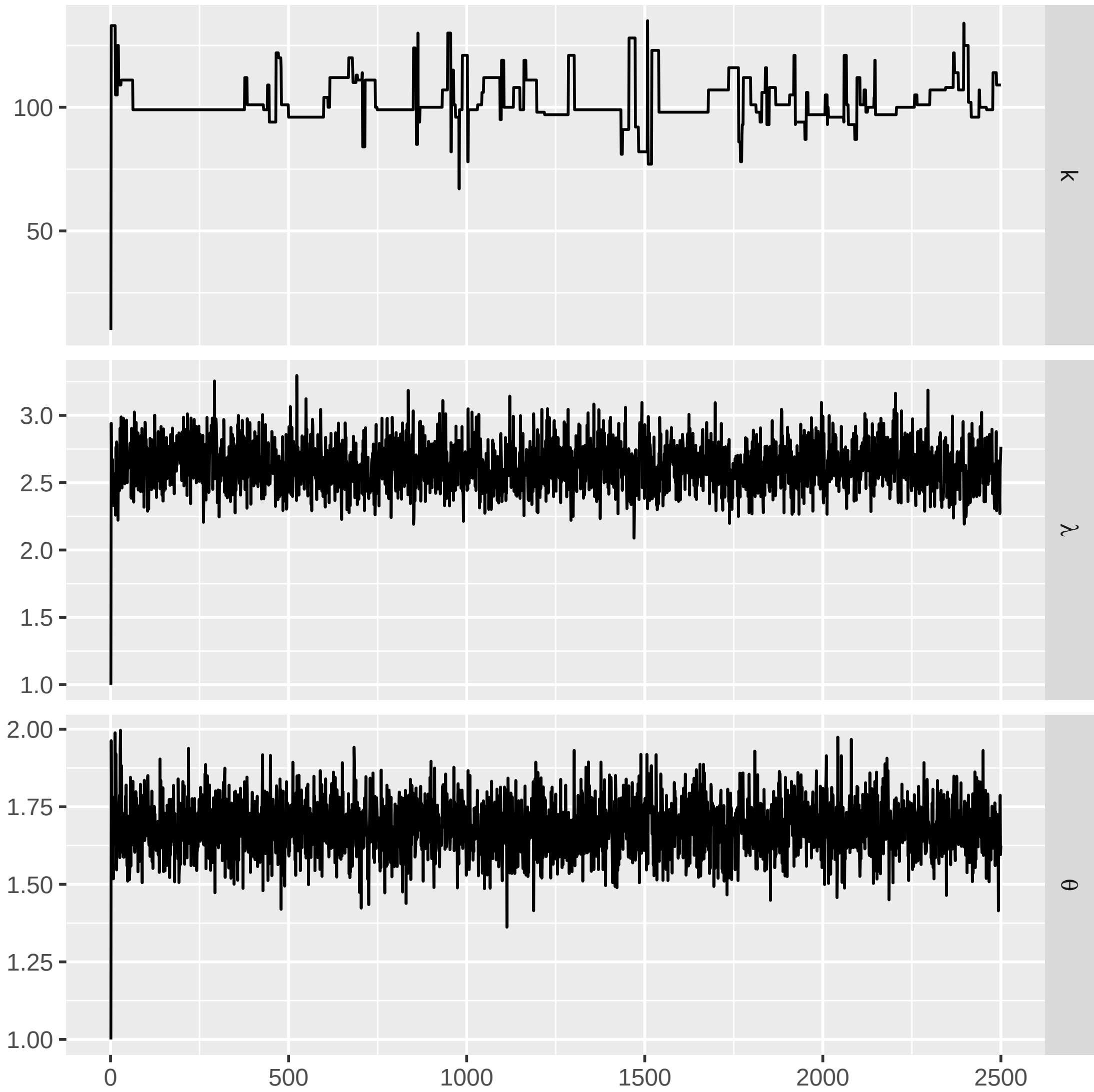

Figure 1: Traceplot of the first 2500 samples for parameters k, λ and θ

Problem 3

A total on N light bulbs are tested to learn about their lifetime. Let the stochastic variable X_i indicate the lifetime of the i^{th} bulb. We assume that the stochastic variables $X_1, \ldots, X_N$ are mutually independent and have an exponential distribution:

$$f(x_i) = \lambda \exp(-\lambda x_i), x_i \geq 0, \quad \text{for } i = 1, \ldots, N$$

The experiment is performed as following: at time $t = 0$ all N bulbs are switched on. At time $t = T$ the experimenter records how many bulbs are still burning and how many have already expired. The observed variables are therefore:

$$y_i = \begin{cases} 0 & \text{if } x_i \leq T \\ 1 & \text{if } x_i > T \end{cases}, \quad i = 1, \ldots, N$$

We are interested in estimating the unknown paramter λ.

Although for this case it is possible to find the maximum likelihood estimator analytically, our aim in this exercise is to use instead the EM algorithm to estimate the parameter λ.

a) Describe briefly the main idea behind the EM-algorithm and in particular what the $Q(\lambda|\lambda^{(k)})$ function is.

b) Derive the $Q(\lambda|\lambda^{(k)})$ function for this problem.

c) Show that the EM algorithm corresponds to the following updating equation for λ:

$$\lambda^{(k+1)} = \frac{N}{K(\lambda^{(k)})}$$

where

$$K(\lambda^{(k)}) = N_0 \left[\frac{1}{\lambda^{(k)}} - T \frac{e^{-\lambda^{(k)}T}}{1 - e^{-\lambda^{(k)}T}} \right] + N_1 \left[\frac{1}{\lambda^{(k)}} + T \right]$$

and N_0 and N_1 indicate the number of observed $y_i = 0$ and $y_i = 1$ respectively.