NTNU

Norwegian University of
Science and Technology

Department of Mathematical Sciences

Examination paper for
TMA4300 Computer Intensive Statistical Methods

Academic contact during examination: Håkon Tjelmeland

Phone: 4822 1896

Examination date: June 6th 2018

Examination time (from–to): 09:00–13:00

Permitted examination support material: C:

- Calculator HP30S, CITIZEN SR-270X, CITIZEN SR-270X College or Casio fx-82ES PLUS with empty memory.
- Statistiske tabeller og formler, Akademika.
- One yellow, stamped A5 sheet with own handwritten formulas and notes.

Other information:

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

Language: English

Number of pages: 4

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 we are only able to sample from the standard uniform distribution Unif(0,1).

a) We want to generate samples from a continuous distribution with density

$$g(x) = \begin{cases} \frac{1}{2}\cos(x) & \text{for } x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right], \\ 0 & \text{otherwise.} \end{cases}$$

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 continuous distribution with density

$$f(x) = \begin{cases} k|x|^\alpha \cos(x) & \text{for } x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right], \\ 0 & \text{otherwise,} \end{cases}$$

where k is a normalising constant and $\alpha \in (0, \infty)$ is a parameter.

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

Problem 2

In this problem we will consider Markov chain Monte Carlo for a toy problem. Let $x \in \{1, 2, 3, 4\}$ be a stochastic variable with distribution $f(x) = x/10, x = 1, 2, 3, 4$. Of course it is easy to sample from this distribution directly, but in this problem we will pretend we do not know how to do this.

To sample realisations from $f(x)$ we will use the Metropolis–Hastings scheme with proposal distribution

$$q(y|x) = \begin{cases} \frac{1}{3} & \text{for } y \neq x, \\ 0 & \text{otherwise} \end{cases}$$

for $x, y \in \{1, 2, 3, 4\}$, where x and y are the current state and the potential new state, respectively.

What is the transition matrix of the Markov chain we are simulating from when using this Metropolis–Hastings algorithm?

Is the Markov chain aperiodic and irreducible? (Remember to give reasons for your answer.)

Problem 3

In this problem we will consider the so called Michaelis-Menten model for enzyme kinetics. In the presence of a catalyst, the model specifies how the chemical reaction rate depends on substrate concentration. Let y denote the chemical reaction rate and let x be the substance concentration. Including an additive error term ε, for $x > 0$ the model specifies

$$y = \frac{\alpha x}{\beta + x} + \varepsilon, \tag{1}$$

where $\alpha, \beta > 0$ are model parameters. Assume we have done n measurements of reaction rate for different substance concentrations. Let $x_1, \ldots, x_n$ denote the substance concentrations for which the measurements are performed and let $y_1, \ldots, y_n$ denote the corresponding measured reaction rates. In the following we will adopt a Bayesian model to analyse the data and the main goal is to estimate the parameters α and β and to predict a new measurement y_0 when the substance concentration is x_0.

We assume the additive error term ε in (1) to have a normal distribution with zero mean and some variance θ, and assume error terms associated to different measurements to be independent. The model has three parameters, θ, α and β, and apriori we assume these to be independent. For the variance θ we assume the improper prior

$$f(\theta) \propto \frac{1}{\theta} \quad \text{for } \theta > 0,$$

and to each of α and β we assign (improper) uniform distributions on $(0, \infty)$.

a) Write an expression for the resulting posterior distribution. It is sufficient to find an expression that is proportional to the posterior distribution.

Derive the full conditional distributions for each of θ, α and β. If possible, specify what parametric family each full conditional belongs to and specify the parameter values.

To explore the posterior distribution we want to define a single-site Metropolis–Hastings algorithm that can simulate from this distribution.

b) For each of θ, α and β specify what proposal distribution you want to use and develop formulas for the corresponding acceptance probabilities. Note that the expressions for the acceptance probabilities should be simplified as much as possible. If your proposal distributions may generate a negative value for θ, α or β your formulas for the acceptance probabilities should take this into account.

Assume you have run your Metropolis–Hastings algorithm for the posterior distribution for M iterations

c) Specify how you would use the Metropolis–Hastings output to estimate the posterior mean values for α and β. Define necessary notation to make your answer precise.

Specify also how you would use the Metropolis–Hastings output to estimate a 90% prediction interval for a new measurement of reaction rate y_0 for substance concentration x_0. Again define necessary notation to make your answer precise.

Problem 4

Let $x_1, \ldots, x_n$ be an observed random sample from a distribution F and let $\mu = \mathrm{E}_F[x]$ and $\sigma^2 = \mathrm{Var}_F[x]$ be the mean value and variance, respectively, in the distribution F.

a) How is the empirical distribution defined in this situation? In general, how is the plug-in estimator for a parameter θ defined? Introduce necessary notation to make your answers precise.

Find the plug-in estimators for μ and σ^2. Simplify the expressions as much as possible.

In the following we use $\widehat{\mu} = s(x) = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ as an estimator for μ. Of course we know that $\widehat{\mu}$ is an unbiased estimator for μ, but in the following we should ignore that we know this fact and use bootstrapping to estimate the bias of $\widehat{\mu}$.

b) Define the ideal bootstrap estimate for the bias of $\widehat{\mu}$. Introduce necessary notation to make your answer precise.

Develop a simple to compute analytical formula for the ideal bootstrap estimator for the bias of $\widehat{\mu}$.

Problem 5

Let $x_1, \ldots, x_n$ be a random sample from a normal distribution with mean value μ and standard deviation σ. Assume, however, that we do not observe the values of $x_1, \ldots, x_n$, we only observe the values rounded down to the nearest integer. Thus, we observe $z_1, \ldots, z_n$ where $z_i = k$ when $x_i \in [k, k+1)$ for $k = 0, \pm 1, \pm 2, \ldots$.

Using observed values $z_1, \ldots, z_n$ we now want to use the EM algorithm to find the maximum likelihood estimates for the parameters μ and σ. When developing formulas for doing this in the problems below you can assume that you have available a function that evaluates the cumulative distribution function $\Phi(x)$ of a standard normal distribution for any value of x, and that you also have available functions that for any values of μ, σ, a and b evaluate the two integrals

$$A(\mu, \sigma, a, b) = \int_a^b x\varphi\left(\frac{x-\mu}{\sigma}\right) \mathrm{d}x \quad \text{and} \quad B(\mu, \sigma, a, b) = \int_a^b x^2\varphi\left(\frac{x-\mu}{\sigma}\right) \mathrm{d}x$$

where $\varphi(x)$ is the density function of a standard normal distribution. *Comment: It is analytically possible to express $A(\mu, \sigma, a, b)$ and $B(\mu, \sigma, a, b)$ in terms of $\varphi(x)$ and $\Phi(x)$, but you do not need to do this.*

a) Setting $x = (x_1, \ldots, x_n)$ and $z = (z_1, \ldots, z_n)$, and letting $f(x; \mu, \sigma)$ denote the joint density of $x_1, \ldots, x_n$, show that $\mathrm{E}\Big[\ln f(x; \mu, \sigma) | z, \mu^{(t)}, \sigma^{(t)}\Big]$ can be expressed as

$$\mathrm{E}\Big[\ln f(x; \mu, \sigma) | z, \mu^{(t)}, \sigma^{(t)}\Big] = -\frac{n}{2}\ln(2\pi) - n\ln\sigma$$
$$-\frac{1}{2\sigma^2}\left(n\mu^2 - 2\mu\alpha(z, \mu^{(t)}, \sigma^{(t)}) + \beta(z, \mu^{(t)}, \sigma^{(t)})\right),$$

and thereby find expressions for $\alpha(z, \mu^{(t)}, \sigma^{(t)})$ and $\beta(z, \mu^{(t)}, \sigma^{(t)})$.

b) Use the EM algorithm setup to develop recursive formulas that can be used to compute the maximum likelihood estimates for μ and σ.

Write also pseudo-code for how we can use bootstrapping to estimate the standard deviations of the maximum likelihood estimators for μ and σ.