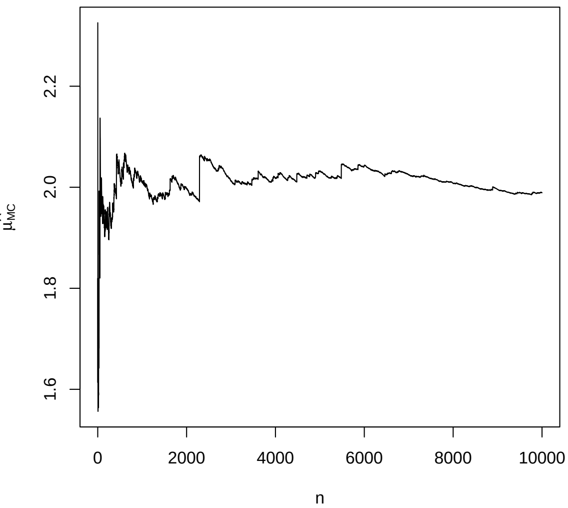

Figure 1: Behaviour of $\hat{\mu}_{MC}$ for increasing n for the model in problem 1

Problem 1

Suppose X, Y are continuous random variables with joint density

$$f_{X,Y}(x, y) \propto \begin{cases} x^2 e^{-xy} & \text{for } 0 < x < 1, y > 0 \\ 0 & \text{elsewhere.} \end{cases}$$

a) Find a method for simulating realisations of X from its marginal density $f_X(x)$.

b) Building on the method in point a), find a method for simulating realistations of X, Y from their joint density. Also find a method for simulating realisations of Y from its marginal density.

c) Derive the expected value of Y using the law of total expectation $E(Y) = E(E(Y|X))$. Suppose that we attempted to estimate $\mu = E(Y)$ using Monte-Carlo integration based on n iid simulated realisations $Y_1, Y_2, \ldots, Y_n$. The behaviour of the Monte-Carlo estimator $\hat{\mu}_{MC}$ for increasing n is shown in Fig. 1. As $n \to \infty$, does $\hat{\mu}_{MC} \to E(Y)$ in mean square, in probabililty, and almost surely, respectively? Would the usual method for estimating the standard error of $\hat{\mu}_{MC}$ be applicable in the present case?

Problem 2

Suppose we observe right sensored survival times of $n = 20$ cancer patients as shown below. Here, if the censoring indicator $\delta_i = 1$, y_i is an observed survival time t_i (in number of years from diagnosis), whereas if $\delta_i = 0$, we instead have a right censored observation meaning that the patient was still alive at time y_i, that is, we know that the event $T_i > y_i$ occured. In total we have 13 right censored observations and 7 uncensored survival times.

```
 i    y delta
 1 1.75     0
 2 2.25     0
 3 0.49     0
 4 1.51     0
 5 1.48     0
 6 2.19     0
 7 1.09     0
 8 0.17     0
 9 1.02     0
10 0.72     0
11 0.18     0
12 0.95     0
13 0.75     0
14 1.16     1
15 1.70     1
16 1.98     1
17 0.94     1
18 3.70     1
19 1.48     1
20 0.82     1
```

We will assume that the survival times $T_1, T_2, \ldots, T_n$ are iid Rayleigh distributed with conditional cumulative distribution function

$$P(T_i \leq t|a) = 1 - e^{-at^2}$$

where $a > 0$ is an unknown parameter.

To represent our prior beliefs about a we will use a lognormal prior

$$\pi(a) = \frac{1}{\sqrt{2\pi\sigma_0 a}} e^{-\frac{(\ln a - \mu_0)^2}{2\sigma_0^2}}$$

where μ_0 and σ_0 are constants.

a) Show the the expected value of the survival times given the parameter a is

$$\mu_T = E(T|a) = \frac{1}{2}\sqrt{\frac{\pi}{a}}.$$

What is the prior on μ_T implied by our choice of prior on a? Based on previous studies, some experts studying this cancer form believe that μ_T is less 1 year with a probability of 5% and above 10 years with a probability also of 5%. How would you need to set μ_0 and σ_0 such that the prior for a is consistent with these beliefs?

Our aim is to use Gibbs sampling to sample from the joint posterior distribution of a and the right censored survival times T_i, $i = 1, 2, \ldots, m$, $m = 13$.

b) What is the full joint conditional of $t_1, t_2, \ldots, t_m$? Find a method for simulating from this full conditional (as a single block).

c) Suppose we want to update a in a separate Metropolis-within-Gibbs step. First write down the full conditional of a, that is, the density of a given $\mathbf{t} = (t_2, t_2, \ldots, t_n)$ up to a normalising constant. You will notice that some factors of the expression are proportional to a Gamma density with a certain shape and rate parameter. Write down the complete expression for $Q(a'|a, \mathbf{t})$ if using this Gamma density as proposal and derive a fully simplified expression for the resulting logarithm of the acceptance probability. Under what circumstances would you expect this to be a good proposal density?

d) Fig. 2 shows trace plots and kernel density estimates of a and t_1 and Fig. 3 shows MCMC samples from the joint posterior of a and $\sum_{i=1}^{n} t_i^2$ based on an implementation of the above Gibbs sampler. Comment on what you see and explain reasons for the somewhat slow convergence of the chain. Briefly discuss ways in which a more efficient single block proposal can be constructed, perhaps building on point b). You do not need to provide details beyond a specifications of the block proposal including any tuning parameters if needed.

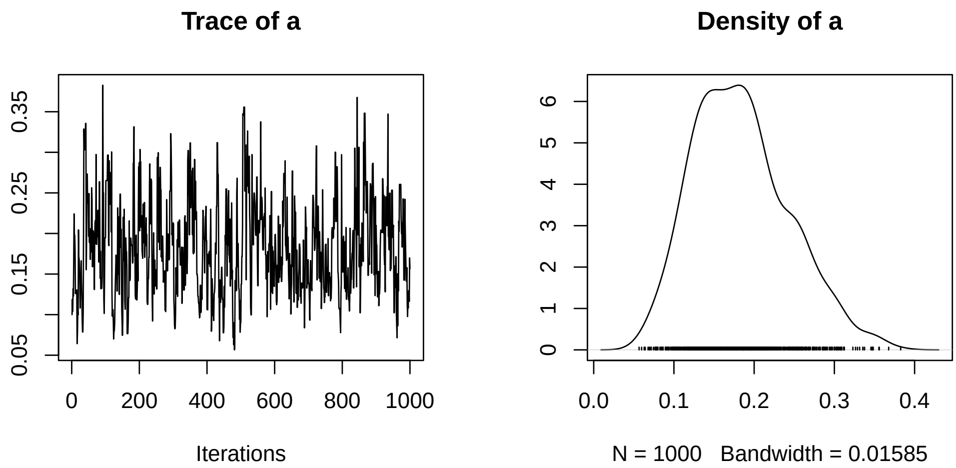

Figure 2: Trace plots and kernel density estimates of a and t_1.

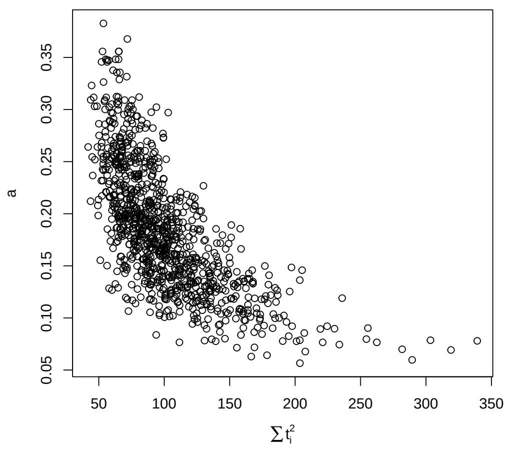

Figure 3: MCMC samples from the joint posterior of a and $\sum_{i=1}^{n} t_i^2$

Problem 3

Suppose that we observe an iid sample from a shifted exponential distribution the density

$$F(x;\lambda,\theta) = \begin{cases} 1 - e^{-\lambda(x-\theta)} & \text{for } x > \theta \\ 0 & \text{for } x \leq \theta \end{cases}$$

where θ and $\lambda > 0$ are unknown parameters. It follows that the maximum likelihood estimators of λ and θ are

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n}(X_i - X_{(1)})}$$

and

$$\hat{\theta} = X_{(1)}$$

where $X_{(1)} = \min(X_1, X_2, \ldots, X_n)$.

In the following suppose that B bootstrap replicates $\hat{\lambda}^b, \hat{\theta}^b$, $b = 1, 2, \ldots, B$ of $\hat{\lambda}, \hat{\theta}$ are available, generated using parameteric bootstrapping from $F(x;\hat{\lambda},\hat{\theta})$.

a) Show that cdf of $X_{(1)}$ is

$$F_{X_{(1)}}(x) = 1 - e^{-n\lambda(x-\theta)}.$$

for $x > \theta$.

How would the bootstrap replicates $\hat{\theta}^b$ of $\hat{\theta}$ be distributed (conditional on $\hat{\lambda}, \hat{\theta}$)? If we neglect Monte-Carlo error in the estimates of the 0.025 and 0.975 quantiles of the boostrap replicates, derive the confidence limits of a 95% confidence interval based on the percentile method.

Under what assumptions are bootstrap confidence intervals based on the percentile method valid? Are these assumptions satisfied in the present case?

What is the probability that the percentile bootstrap interval contains θ in the present case?

b) Explain in detail how an alternative bootstrap t confidence interval for θ can be computed. Derive an expression for the resulting lower and upper confidence limit in terms of observable quantities. Would such an interval be exact?

Suppose that we estimate the expected value of $\hat{\theta}$ from the bootstrap samples by

$$\widehat{E(\hat{\theta})} = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}^b$$

and its bias, $\text{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta$, by

$$\widehat{\text{Bias}(\hat{\theta})} = \frac{1}{B}\sum_{b=1}^{B} \hat{\theta}^b - \hat{\theta}$$

Based on this we constuct a bias-corrected estimator

$$\hat{\theta}_c = \hat{\theta} - \widehat{\text{Bias}(\hat{\theta})}.$$

c) Use the cdf in point a) to show that

$$E(\hat{\theta}) = \theta + \frac{1}{n\lambda}.$$

Can you make a similar statement about $E(\hat{\theta}^b|\hat{\lambda}, \hat{\theta})$?

Will the bias-corrected estimator $\hat{\theta}_c$ be unbiased for θ? Hints: According to the law of total expectation $E(\hat{\theta}_c) = EE(\hat{\theta}_c|\hat{\theta}, \hat{\lambda})$. You may also need to consider the memoryless property of the exponential distribution.