MATH 80601A *Bayesian modelling*

Practice midterm examination

Exam booklet

Examiner: Léo Belzile,

Instructions: The time allotted for the examination is 180 minutes. You may answer in either English or French. No written material may be brought into the examination.

There are a total of 21 marks available in the exam paper, the distribution of which can be found in the right margin in brackets.

You must hand back the **exam booklet** at the end of the examination.

Crib sheet: A Poisson random variable with mean $\lambda > 0$ denoted $\mathsf{Poisson}(\lambda)$, has density

$$f(x) = \frac{\lambda^x}{x!}\exp(-\lambda), \quad x \in 0, 1, \ldots$$

A Gaussian random variable with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$, denoted $\mathsf{Gauss}(\mu, \sigma^2)$, has density

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}, \qquad x \in \mathbb{R}.$$

An inverse gamma random variable with shape $\alpha > 0$ and rate $\beta > 0$, denoted $\mathsf{inv.gamma}(\alpha, \beta)$, has density

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{-\alpha-1}\exp(-\beta/x), \qquad x \in \mathbb{R}.$$

where $\Gamma(\cdot)$ is the gamma function. The latter satisfies $\Gamma(a+1) = a\Gamma(a)$ for $a \geq 0$.

Question:	1	2	Total
Points:	5	16	21
Score:			

Question 1. Short answer questions

1.1 If X is a continuous random variable with density function $f(x)$, is $f(x_0)$ a probability for given x_0? Justify your answer. [1 pt]

1.2 **Inversion sampling**: for an absolutely continuous distribution function F with quantile function F^{-1} and $U \sim \mathsf{unif}(0,1)$ uniform, prove that $F^{-1}(U)$ has distribution function F. [2 pts]

1.3 Give an example of a distribution $p(Y \mid \theta)$ which may admit as conjugate prior for θ [2 pts]

  i. the gamma distribution,

  ii. the beta distribution.

Question 2. Modelling unemployment rates

We consider a times series of $n = 142$ monthly harmonised unemployment rates (in percentage) of women in Norway for the period running from April 2000 to January 2012. We treat the sample $\boldsymbol{y}_{1:n} = (y_1, \ldots, y_n)^\top$ as the realization of an autoregressive process of order 1, denoted AR(1), of the form

$$Y_t \mid Y_{t-1} = y_{t-1} \sim \mathsf{Gauss}(\alpha + \phi y_{t-1}, \nu), \qquad t = 1, \ldots, n; \tag{1}$$

where α is a location parameter, ν is the conditional **variance** and ϕ is the autoregressive parameter. Throughout, we perform inference conditional on the observed initial value y_0 for March 2000, treating the latter as a fixed quantity. Recall that an AR(1) process is stationary if $|\phi| < 1$, with the unconditional variance of $\sigma^2/(1-\phi^2)$.

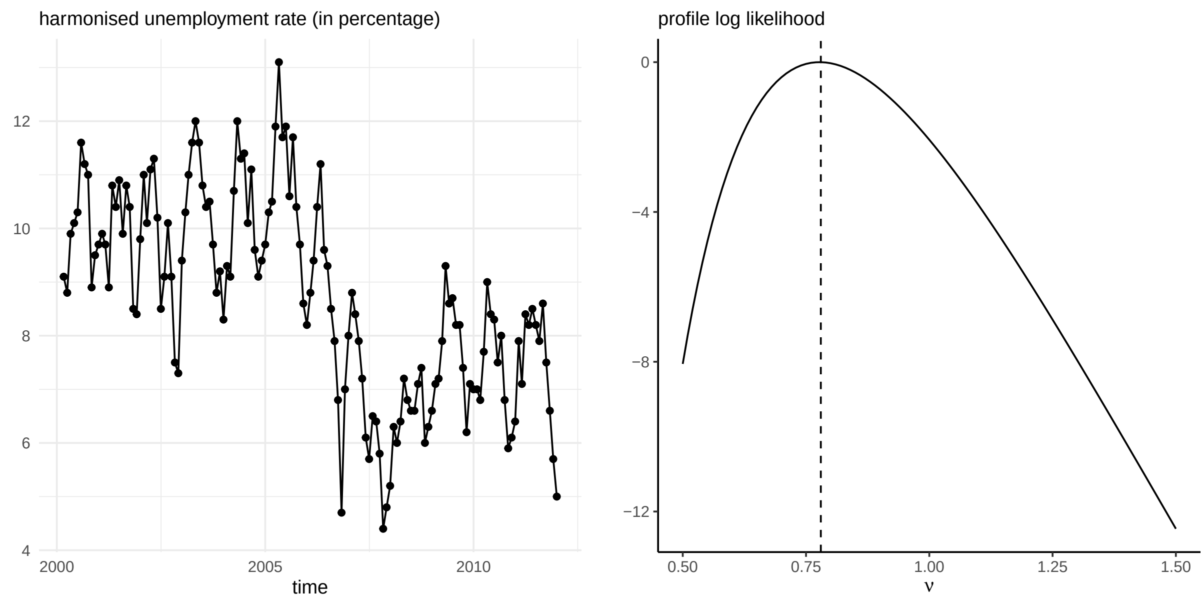

Figure 1: Left: time series of monthly women harmonised unemployment rate over time. Right: Profile log likelihood for the variance parameter ν of eq. (1). The curve has been shifted down so that the profile log likelihood is zero at the maximum likelihood estimate. The vertical dashed line indicates the maximum likelihood estimate.

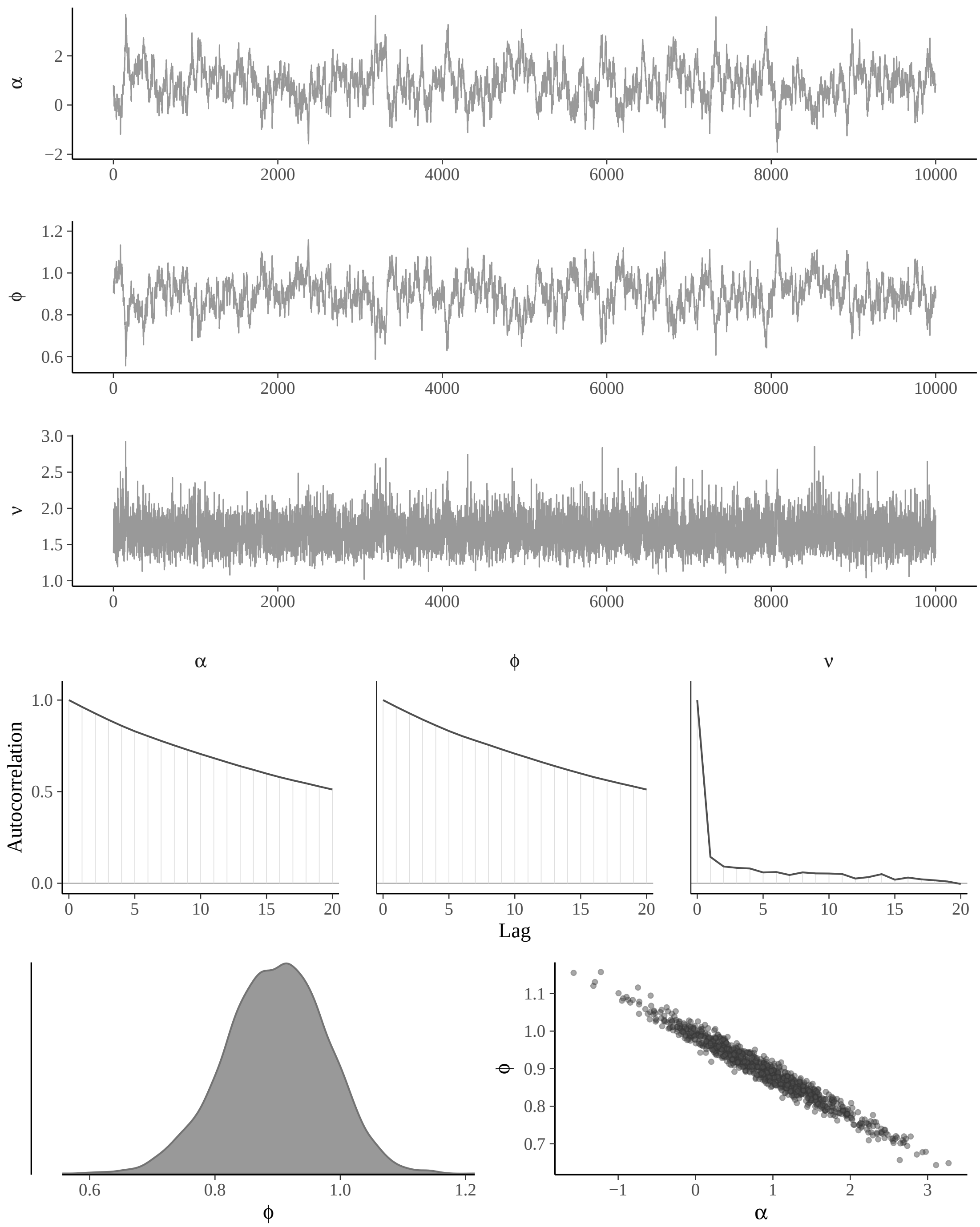

Figure 2: Diagnostics of the Markov chains obtained by running a Gibbs sampler to obtain 10 000 posterior draws from α, ϕ, ν, with trace plots (top), correlograms of individual chains (middle), histogram of the autocorrelation coefficient (bottom left) and scatterplot of pairs (α, ϕ) (bottom right). The estimated effective sample sizes are $\mathsf{ESS}(\alpha) = 226$, $\mathsf{ESS}(\phi) = 217$ and $\mathsf{ESS}(\nu) = 3644$.

2.1 Write down the joint likelihood $L(\phi, \alpha, \nu; \boldsymbol{y}_{1:n})$. [2 pts]

2.2 Consider the reference prior [1 pt]

$$p(\alpha, \phi, \nu) \propto \nu^{-1}\, \mathrm{I}(\nu > 0)\mathrm{I}(\alpha \in \mathbb{R})\mathrm{I}(\phi \in \mathbb{R}). \tag{2}$$

Is this prior proper? Justify your answer.

2.3 Using the reference prior of Equation (2), derive and provide explicit expressions for the parameters of the posterior distributions [4 pts]

$$\begin{gathered} p(\alpha \mid \phi, \nu, \boldsymbol{y}_{1:n}) \\ p(\phi \mid \alpha, \nu, \boldsymbol{y}_{1:n}) \\ p(\nu \mid \alpha, \phi, \boldsymbol{y}_{1:n}) \end{gathered}$$

Explain how you would use these to obtain posterior draws using a Gibbs sampler.

2.4 Suppose you to select a more informative prior for the variance by taking $\nu \sim \mathsf{inv.gamma}(\beta, \beta)$. Derive the expected value of the $\mathsf{inv.gamma}(\beta, \beta)$ prior and use moment matching to select the value β such that the apriori expected value for ν is 1.5. [2 pts]

2.5 Figure 2 on page 3 shows diagnostic plots for the Gibbs sampler. Comment on the stationarity and mixing, and diagnose any problems related to sampling. [2 pts]

2.6 Figure 2 shows some posterior draws of ϕ exceed one, which corresponds to AR(1) nonstationary models with infinite variance. How could you enforce stationarity? [1 pt]

2.7 In line with the previous two subquestions, describe modifications of your algorithm that would improve the efficiency of the sampling of posterior draws. [2 pts]

2.8 For February 2012, the analysis yields a one-step ahead forecast, obtained from the posterior predictive distribution $p(\widetilde{y}_{n+1} \mid \boldsymbol{y}_{1:n})$, of 5.39% with an equitailed 50% credible interval of $[4.48\%, 6.30\%]$. Explain in your own words how you would calculate such interval based on the posterior draws $(\alpha_{(b)}, \phi_{(b)}, \sigma_{(b)})$ $(b = 1, \ldots, B)$. [2 pts]