

# BCCS 2008/09: Graphical models and complex stochastic systems: Take-home open-book test

1. In not more than 4 or 5 sentences overall, make brief notes on some of the advantages of using conditional independence and graphical modelling ideas in formulating statistical models, and carrying out certain computations using these models.
2. You are given  $n$  independent observations  $x_1, x_2, \dots, x_n$  that can be taken to be drawn from Poisson distributions,

$$p(x_i|\theta) = e^{-\theta t_i} (\theta t_i)^{x_i} / x_i!$$

where  $\theta > 0$  is an unknown parameter, and the  $\{t_i\}$  are known positive constants.

- (a) Write down the likelihood for  $\theta$  and the logarithm of the likelihood. Hence find the maximum likelihood estimate of  $\theta$  in this situation. (You should verify that this value does indeed give a maximum of the likelihood).
  - (b) Now assume that you have prior information about  $\theta$ , that is expressed by a Gamma( $\alpha, \beta$ ) distribution, that is, it has probability density  $p(\theta|\alpha, \beta) = \beta^\alpha \theta^{\alpha-1} \exp(-\beta\theta) / \Gamma(\alpha)$  for  $\theta > 0$ , 0 otherwise. [ $\Gamma(\alpha)$  is the Gamma function, defined as  $\int_0^\infty x^{\alpha-1} e^{-x} dx$ .] You may assume that the expectation of this prior distribution is  $\alpha/\beta$ . Find the posterior distribution of  $\theta$  given the  $n$  observations above. Write down the posterior expectation of  $\theta$ , and comment on its relationship to the maximum likelihood estimate you found in (a).
3. Consider a hidden Markov model using the notation of Section 7.6 of the notes.
    - (a) Show that the function  $r_t(x_t)$  can be interpreted as  $p(x_t, y_{\leq t})$  and hence verify the final assertion of the section, namely that

$$p(x_t|y_{\leq t}) = \frac{r_t(x_t)}{\sum_{x_t} r_t(x_t)}.$$

- (b) Explain why it is true that

$$p(x_t, y_{\leq t+1}) = p(x_t, y_{\leq t}) \sum_{x_{t+1}} p(x_{t+1}|x_t) p(y_{t+1}|x_{t+1}) = r_t(x_t) \sum_{x_{t+1}} g_{t+1}(x_t, x_{t+1})$$

and hence suggest an efficient algorithm for computing the ‘one-step behind’ smoothing distributions  $p(x_t|y_{\leq t+1})$  for all  $t$ .

- (c) Show that

$$p(x_{t+1}|y_{\leq t}) = \sum_{x_t} p(x_t|y_{\leq t}) p(x_{t+1}|x_t)$$

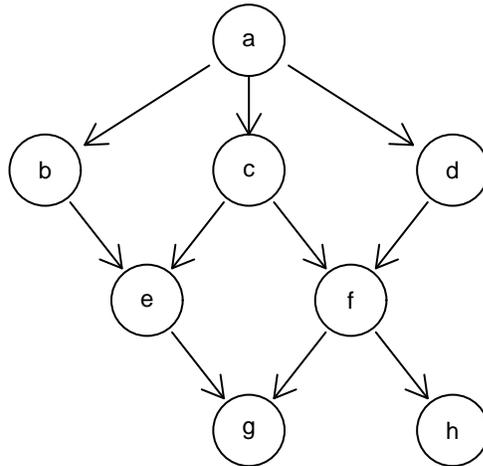
and hence suggest an efficient algorithm for computing the ‘one-step ahead’ predictive distributions  $p(x_{t+1}|y_{\leq t})$  for all  $t$ .

4. Consider the following model for discrete waiting-time data  $y_1, y_2, \dots, y_n$ . Given parameters  $\{\theta_i\}$ , the  $y_i$  are modelled as independent Negative Binomial variables, with

$$p(y_i|\theta_i) = \frac{(y_i + r - 1)!}{y_i!(r - 1)!} \theta_i^r (1 - \theta_i)^{y_i} \quad y_i = 0, 1, 2, \dots$$

where  $r$  is fixed at the value 2. In turn the  $\{\theta_i\}$  are drawn (conditionally) independently from a Beta prior  $p(\theta_i|\alpha, \beta) = [\Gamma(\alpha + \beta) / \{\Gamma(\alpha)\Gamma(\beta)\}] \theta_i^{\alpha-1} (1 - \theta_i)^{\beta-1}$  for  $0 < \theta_i < 1$ , and  $\alpha$  and  $\beta$  are independent Exponential random variables with rate 1:  $p(\alpha) = e^{-\alpha}$ ,  $p(\beta) = e^{-\beta}$  for  $\alpha, \beta > 0$ .

- (a) Draw a DAG representing this model.
- (b) Write down the joint distribution of all random variables  $(\alpha, \beta, \{\theta_i\}, \{y_i\})$ .
- (c) Show that the full conditional distribution of  $\theta_i$  in this model is a Beta distribution, and find its parameters.
- (d) Write down an expression giving the full conditional distribution of  $\alpha$  in this model, up to a constant of proportionality; simplify it as far as possible. (*You may assume that the full conditional for  $\beta$  is somewhat similar in form.*) Assuming that you do not recognise this expression as the density function for a standard distribution, explain why this poses a problem for Gibbs sampling in this model. Without going into any details, is there any other MCMC method that might be usable in this context?
- (e) **Bonus question.** You can set up this model, with a small set of data, in Winbugs, using the files that can be downloaded from the course website. Run this code to make inference about the parameters, and write brief comments on the results.
5. The directed acyclic graph below expresses the relationship between 8 variables in a simple graphical model.



Decide which of the following conditional independence statements can be inferred from the graph.

- (a)  $e \perp\!\!\!\perp (a, d, f) \mid (b, c)$
- (b)  $b \perp\!\!\!\perp h \mid d$
- (c)  $a \perp\!\!\!\perp (e, g) \mid (b, c, d)$
- (d)  $a \perp\!\!\!\perp (g, h) \mid (c, d, e)$

In the case of (d), verify your answer *both* using a graphical argument, and by writing down an expression for  $p(a, c, d, e, g, h)$  and showing that it factorises appropriately.