Probabilistic Modelling and Reasoning, Tutorial Question Sheet 3 (for Week 5)

School of Informatics, University of Edinburgh

Instructor: Amos Storkey

Prob. Sheet. ID: 11922.1.1

1. Consider the Bernoulli random variable X with mean θ, and suppose we have observed n_h occurrences of $X = 1$ and n_t occurrences of $X = 0$. The prior distribution for $\theta \sim \text{Beta}(\alpha_h, \alpha_t)$. Show that the posterior mean value of θ lies between the prior mean and the maximum likelihood estimate of θ. To do this, show that the posterior mean can be written as λ times the prior mean plus $(1-\lambda)$ times the maximum likelihood estimate, for $0 \leq \lambda \leq 1$. HINT: to make this easier, verify that the required value of $\lambda = \alpha/(n+\alpha)$, where $\alpha = \alpha_h + \alpha_t$ and $n = n_h + n_t$. This illustrates the concept of the posterior distribution being a compromise between the prior distribution and the maximum likelihood solution. [Bishop qu 2.7].

2. Consider the beta distribution $p(\theta) = c(\alpha, \beta)\theta^{\alpha-1}(1-\theta)^{\beta-1}$, where $c(\alpha, \beta)$ is a normalizing constant. The mean of this distribution is $E[\theta] = \alpha/(\alpha+\beta)$. For $\alpha, \beta > 1$ the distribution is unimodal (i.e. it has only one maximum). Find the value θ^* where this maximum is attained, and compare it to the mean. For what values of α and β to the mean and θ^* coincide?

3. Consider a crowdsourcing situation where workers are provided with a series of images. X denotes whether each image has a car in it or not ($X = 1$ is a car, $X = 2$ is not). In general this is not known, but we do know that 30% of the images have cars in. Supervisors get workers to state if there is a car in the image or not and report it using C. $C = 1$ means a car, and $C = 2$ is not a car. People make mistakes by failing to spot cars in an image 10% of the time, and recognising cars when there are none 1% of the time.

   However there are four supervisors, and one of these supervisors is under a misunderstanding and instructed workers incorrectly, asking people to label the $C = 2$ for car and $C = 1$ for no car. Your training data has come from the results of this crowdsourcing exercise from one of the supervisors, but you do not know which (it is equally likely to be any of them). The training data takes the form $C = 1, 1, 2$ for each of three different images presented to three different crowdsourcing workers. All the training data is from workers managed by the same supervisor.

   You will also receive one more label shortly (from a worker supervised by the same supervisor).

   Draw the belief network for this system using a plate notation (but separating out the future case from those in the plate). Attach the relevant probabilities to the network. What do you consider to be parameters, what do you consider to be visible variables, and what do you consider to be latent variables in this model?

   What do you estimate for the probability of the next label you receive (from a worker supervised by the same supervisor) being $C = 1$?

   You now receive the further result ($C = 1$) from a worker working with the same supervisor. What do you estimate for the probability of X being 1 for this case?