FINAL EXAM

ISyE 8843: Bayesian Statistics
Brani Vidakovic
due before noon, Thursday 12/9/2004.

Name _____

1. Bayesian Wavelet Shrinkage. This open ended question is essentially asking to select a data set with noise present in it (a noisy signal, function, or noisy image), transform the data to the wavelet domain, apply shrinkage by suitably developed Bayes rules on wavelet coefficients, and back-transform shrunk coefficients (alias Bayes estimates) to the domain of original data.

Recent Tech Report 34/2004 at `http://www.isye.gatech.edu/~brani/isyestat/` updates the Handout 21 with some recent references, and you may use some of the procedures supplied there. However, the question is open ended and you may propose your own method and use the software of your choice, even BUGS.

Illustration. If you are out of idea what Bayes model to use to shrink in the wavelet domain, here is an example of a possible solution: Prove that for $[d|\theta] \sim \mathcal{N}(\theta, 1)$, $[\theta|\tau^2] \sim \mathcal{N}(0, \tau^2)$, and $[\tau^2] \sim (\tau^2)^{-3/4}$ the posterior is unimodal if $0 < d^2 < 2$ and bimodal otherwise with the second mode

$$\delta(d) = \left(1 - \frac{1 - \sqrt{1 - 2/d^2}}{2}\right) d.$$

Generalize to $[d|\theta] \sim \mathcal{N}(\theta, \sigma^2)$, σ^2 known, and apply **the largest mode shrinkage**. Is this shrinkage of thresholding type?

Use approximation $(1 - u)^\alpha \sim (1 - \alpha u)$ for u small to argue that the largest mode shrinkage is close to a James-Stein-type rule $\delta^*(d) = \left(1 - \frac{1}{2d^2}\right)_+ d$, where $(f)_+ = \max\{0, f\}$.

2. Alarm. Refer to the `alarm` example. Find $P(J1, M1|B1)$ i.e., the probability that both John and Mary call, given the burglary, in three ways:
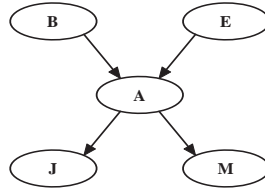

Figure 1: Alarm Bayes Network

(i) exact, by calculating

$$P(J1, M1|B1) = \frac{1}{P(B1)} P(B1, J1, M1) = \frac{1}{P(B1)} \sum_{E,A} P(B1, E, A, J1, M1) = \ldots$$

(ii) using Kevin Murphy's BNT, and
(iii) using BUGS.
Note that $P(J1, M1|B1) = P(J1|M1, B1) \cdot P(M1|B1)$.
The `alarm` conditional probabilities are:

B0	B1
0.999	0.001

E0	E1
0.998	0.002

A0	A1	condition
0.999	0.001	B0 E0
0.71	0.29	B0 E1
0.06	0.94	B1 E0
0.05	0.95	B1 E1

J0	J1	condition
0.95	0.05	A0
0.10	0.90	A1

M0	M1	condition
0.99	0.01	A0
0.30	0.70	A1

3. Change Point Analysis. The time intervals in days between successive coal mining accidents involving more than ten men killed during the period 1851 to 1962 in Great Britain were reported by Jarrett (1979). This dataset given in `http://www.isye.gatech.edu/~brani/isyebayes/data/mining_acc.dat` is a corrected and extended version of the dataset previously given by Maguire *et al.* (1952). The number of accidents occurring over a given time period can be looked on as having a Poisson distribution. For this problem Jarrett/Maguire's data are binned in time intervals of 1 year and given below as a MATLAB input.

```
tt=[1851:1962];
minedata = [4,5,4,1,0,4,3,4,0,6,3,3,4,0,2,6,3,3,5,4,5,3,1,4,4,1,5,5,3,...
            4,2,5,2,2,3,4,2,1,3,2,2,1,1,1,1,3,0,0,1,0,1,1,0,0,3,1,0,3,2,...
            2,0,1,1,1,0,1,0,1,0,0,0,2,1,0,0,0,1,1,0,2,3,3,1,1,2,1,1,1,1,...
            2,4,2,0,0,0,1,4,0,0,0,1,0,0,0,0,0,1,0,0,1,0,1];
```

• If you plot `minedata` $(Y_i, i = 1, \ldots, 112)$ against `tt` (years 1851-1962), you will notice that there is a change point located somewhere around year 1890.

• Assume the model

$$
\begin{aligned}
{[Y_i|\tau, \lambda]} &\sim \mathcal{P}oi(\lambda), \ i = 1, \ldots, \tau \\
{[Y_i|\tau, \mu]} &\sim \mathcal{P}oi(\mu), \ i = \tau + 1, \ldots, n \\
{[\tau]} &\sim \mathcal{DU}(n), \ \text{ i.e., } \ P(\tau = k) = \frac{1}{n}, \ 1 \le k \le n \\
{[\lambda]} &\sim \mathcal{G}amma(\alpha_\lambda, \beta_\lambda) \\
{[\mu]} &\sim \mathcal{G}amma(\alpha_\mu, \beta_\mu)
\end{aligned}
$$

Here $n, \alpha_\lambda, \beta_\lambda, \alpha_\mu,$ and β_μ are known. For a more complex formulation (hyperprior on the scale of Gamma's) see Robert and Casella (1999), pages 432–434, also page 444. If this problem intrigues you beyond the requirements for this final, you may want to read Carlin *et al.* (1992).

• Exact posteriors are intractable due to τ's contribution to the model, but full conditionals are straightforward. Show that the full conditionals are

$$
\begin{aligned}
P(\tau = k|\lambda, \mu, Y) &\propto \lambda^{\alpha_\lambda + \sum_{i=1}^{k} Y_i - 1} \times e^{-(\beta_\lambda + k)\lambda} \times \mu^{\alpha_\mu + \sum_{i=k+1}^{n} Y_i - 1} \times e^{-(\beta_\mu + (n-k))\mu}, \ \ 1 \le k \le n \\
{[\lambda|\tau, \mu, Y]} &\sim \mathcal{G}amma\left(\alpha_\lambda + \sum_{i=1}^{\tau} Y_i, \beta_\lambda + \tau \right) \\
{[\mu|\tau, \lambda, Y]} &\sim \mathcal{G}amma\left(\alpha_\mu + \sum_{i=\tau+1}^{n} Y_i, \beta_\mu + (n - \tau) \right)
\end{aligned}
$$

To generate τ the probabilities above have to be normalized by

$$
\sum_{k=1}^{n} \left\{ \lambda^{\alpha_\lambda + \sum_{i=1}^{k} Y_i - 1} \times e^{-(\beta_\lambda + k)\lambda} \times \mu^{\alpha_\mu + \sum_{i=k+1}^{n} Y_i - 1} \times e^{-(\beta_\mu + (n-k))\mu} \right\},
$$

where $\sum_{i=n+1}^{n} Y_i$ is defined as 0.

• Implement MCMC. Propose hyperparameters $\alpha_\lambda, \beta_\lambda, \alpha_\mu,$ and β_μ and initial values for λ and μ that work well. Index $i = 1$ corresponds to year 1851, and index $i = n (= 112)$ corresponds to year 1962.

• Discuss the MCMC results for $\tau, \lambda,$ and μ.

Remark: As Abhyuday pointed out, by ignoring constant terms the probability $P(\tau = k|\lambda, \mu, Y)$ simplifies to

$$P(\tau = k|\lambda, \mu, Y) = \frac{e^{-(\lambda-\mu)k} \times (\lambda/\mu)^{\sum_{i=1}^{k} Y_i}}{\sum_{k=1}^{n} \left\{ e^{-(\lambda-\mu)k} \times (\lambda/\mu)^{\sum_{i=1}^{k} Y_i} \right\}}, 1 \leq k \leq n,$$

or

$$P(\tau = k|\lambda, \mu, Y) = \frac{\exp\{ (\mu - \lambda)k + \log(\lambda/\mu) \sum_{i=1}^{k} Y_i\}}{\sum_{k=1}^{n} \exp\{ (\mu - \lambda)k + \log(\lambda/\mu) \sum_{i=1}^{k} Y_i\}}, 1 \leq k \leq n.$$

Much better for MCMC! Ying pointed out the missing λ, μ typo. Thanks!

REFERENCES

Carlin, B. P., Gelfand, A. E., and Smith, A. F. M. (1992). Hierarchical bayesian Analysis of change point problems, *Applied Statistics* (Ser. C), 41, 389–405.

Jarrett, R. G. (1979). A note on the intervals between coal mining disasters. *Biometrika*, 66, 191–193.

Maguire, B. A., Pearson, E.S. and Wynn, A. H. A. (1952). The time intervals between industrial accidents. *Biometrika*, 39, 168–180.

Raftery, A.E., and Akman, V.E. (1986). Bayesian analysis of a Poisson process with a change-point. *Biometrika*, 73, 85–89.

Robert, C. and Casella, G. (1999). *Monte Carlo Statistical Methods*, Springer Verlag, New York