Stat 544 Final Exam

May 1, 2007

I have neither given nor received unauthorized assistance on this examination.

____________________________________ ________________
signature date

1. A classic example in the Bates and Watts non-linear regression text concerns the "Rumford cooling experiment," where after friction heating a cannon, Rumford stopped the heating process and observed the temperature of the cannon over time as it cooled down. For

y_t = the measured temperature t minutes after the end of heating

Rumford's data were

t	4	5	7	12	14	16	20	24	28	31	34	37.5	41
y_t	126	125	123	120	119	118	116	115	114	113	112	111	110

For parameters $\beta_0, \beta_1 > 0, \theta > 0,$ and $\sigma^2 > 0$ consider a "data model" for the y_t of the form

$$y_t = \left(\beta_0 + \beta_1 \exp(-\theta t)\right)(1 + \varepsilon_t) \qquad (*)$$

for ε_t's that are iid $\mathrm{N}(0, \sigma^2)$. (This is a model that might correspond to a situation where the standard deviation of temperature measurement is proportional to the real temperature.)

a) Write out a likelihood function for this model. You may use the notation $f(y \mid \mu, \sigma^2)$ for the univariate normal pdf with mean μ and variance σ^2, and write $\prod_t$ without completely specifying the values of t indicated above.

b) Suppose that one places priors on the variables $\beta_0, \beta_1, \theta,$ and σ^2, and declares them to be *a priori* independent. Make below an appropriate DAG for describing this scenario. You may show only $y_4, y_5,$ and y_{41} (and not the rest of the y_t's) in order to save space.

c) Based on your DAG from b), are $y_4, y_5,$ and y_{41} guaranteed to be conditionally independent given $\beta_0, \beta_1,$ and θ? Say YES or say NO and then explain carefully why your answer is correct.

d) Suppose that the continuous priors one uses for $\beta_0, \beta_1,$ and θ have respective pdfs $g_0(\beta_0), g_1(\beta_1),$ and $\mathrm{g}(\theta)$ and that one uses an improper prior for σ^2 with "density" proportional to $1/\sigma^2$ (and that the prior independence mentioned in b) is used). In a Gibbs sampling algorithm for this model, when one is at $\beta_0^{\text{current}}, \beta_1^{\text{current}}, \theta^{\text{current}},$ and $\sigma^{2\ \text{current}}$ and must update σ^2, what is a function of σ^2 proportional to the pdf from which one must sample? (Make use of the notation of part a).)

On the next page there is some `WinBUGS` output made with a prior for $\beta_0, \beta_1, \theta,$ and σ^2 as described in part d) with g_0 Uniform on $(-300, 300)$, g_1 Uniform on $(0, 400)$, and g Uniform on $(0,1)$. Use it to answer e) and f).

e) According to the model (*), the cannon achieves $e^{-1} = 63\%$ of its drop in temperature $t = 1/\theta$ minutes after it begins cooling. What are 95% credible limits for this time?

f) According to the model (*), the standard deviation of temperature measurement at a given temperature is positively related to temperature. What are 95% credible limits for the standard deviation of measurement with cannon temperature at its steady state value?

Printout for Parts e) and f)

```
model {

  b0 ~ dunif (-300,300)
  b1 ~ dunif (0,400)
  theta ~ dunif (0,1)
  lsigma ~ dflat ()
  taustar <- exp(-2*lsigma)
  sig0 <- 1/sqrt(taustar/pow(b0+b1,2))
  siginfinity <- 1/sqrt(taustar/pow(b0,2))

  for (i in 1:13) {
    mu[i] <- b0+b1*exp(-theta*t[i])
  }

for (i in 1:13) {
    tau[i] <- taustar/pow(mu[i],2)
  }

  for (i in 1:13) {
    y[i] ~ dnorm (mu[i],tau[i])
  }

}

list(y=c(126,125,123,120,119,118,116,115,114,113,112,111,110),t=c(4,5,7,12,14,16,20,24,28,31,34,37.5
,41))
```

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
b0	105.9	1.322	0.03047	102.8	106.0	108.0	4001	318000
b1	23.17	1.016	0.02227	21.58	23.05	25.52	4001	318000
mu[1]	125.6	0.2773	0.003262	125.1	125.6	126.2	4001	318000
mu[2]	124.8	0.2378	0.002274	124.4	124.8	125.3	4001	318000
mu[3]	123.4	0.1834	7.742E-4	123.0	123.4	123.8	4001	318000
mu[4]	120.2	0.1609	0.002139	119.9	120.2	120.5	4001	318000
mu[5]	119.1	0.1683	0.00262	118.8	119.1	119.5	4001	318000
mu[6]	118.1	0.1735	0.002866	117.8	118.1	118.5	4001	318000
mu[7]	116.3	0.1711	0.002768	116.0	116.3	116.7	4001	318000
mu[8]	114.8	0.1548	0.002068	114.5	114.8	115.1	4001	318000
mu[9]	113.5	0.1381	9.703E-4	113.2	113.5	113.8	4001	318000
mu[10]	112.6	0.1373	4.431E-4	112.4	112.6	112.9	4001	318000
mu[11]	111.9	0.1535	0.001417	111.6	111.9	112.2	4001	318000
mu[12]	111.1	0.1922	0.002863	110.7	111.1	111.5	4001	318000
mu[13]	110.4	0.2444	0.00439	110.0	110.4	110.9	4001	318000
sig0	0.412	0.1047	8.206E-4	0.265	0.3934	0.669	4001	318000
siginfinity	0.3379	0.08559	6.369E-4	0.2176	0.3227	0.5477	4001	318000
theta	0.04025	0.004764	1.029E-4	0.03058	0.04025	0.04984	4001	318000

g) The model (*) has non-constant variance for the y_t's. An alternative model (with constant variance) is

$$y_t = \beta_0 + \beta_1 \exp(-\theta t) + \varepsilon_t \qquad (**)$$

where as before the ε_t are iid $\mathrm{N}(0,\sigma^2)$. Say very carefully how you would *compute* a Bayes factor for comparing models (*) and (**) with priors (for both cases) of independence with marginals

$$\beta_0 \sim \mathrm{U}(-300,300), \beta_1 \sim \mathrm{U}(0,400), \theta \sim \mathrm{U}(0,1), \text{ and } 1/\sigma^2 \sim \Gamma(.01,.01)$$

using appropriate simulations.

h) The temperatures on page 2 were pretty clearly measured to the nearest degree. It seems possible that treating them as "infinite number of decimal places" values might be a bad thing to do. An alternative is to treat them as interval censored values. Finish the `WinBUGS` model statement code below necessary to recognize the interval censoring.

```
model {

  b0 ~ dunif (-300,300)
  b1 ~ dunif (0,400)
  theta ~ dunif (0,1)
  lsigma ~ dunif (-10,10)
  taustar <- exp(-2*lsigma)
  sig0 <- 1/sqrt(taustar/pow(b0+b1,2))
  siginfinity <- 1/sqrt(taustar/pow(b0,2))

  for (i in 1:13) {
    mu[i] <- b0+b1*exp(-theta*t[i])
  }

for (i in 1:13) {
    tau[i] <- taustar/pow(mu[i],2)
  }

  for (i in 1:13) {
    low[i] <- y[i]-.5
    high[i] <- y[i]+.5

    z[i] ~                                                      #← finish this

  }

}
```

i) Below are some `WinBUGS` results from running the code from part h). What, if any, important differences do you see between the results here and those on page 4? In this application, how does ignoring the interval censoring affect one's perception of "error variance"?

node	**mean**	**sd**	**MC error**	**2.5%**	**median**	**97.5%**	**start**	**sample**
b0	105.4	0.8359	0.02476	103.7	105.5	107.1	15001	300000
b1	23.52	0.6636	0.01905	22.24	23.44	24.95	15001	300000
mu[1]	125.6	0.141	0.002214	125.3	125.6	125.9	15001	300000
mu[2]	124.8	0.1206	0.001533	124.6	124.8	125.1	15001	300000
mu[3]	123.4	0.09647	9.023E-4	123.2	123.4	123.6	15001	300000
mu[4]	120.3	0.09945	0.00216	120.1	120.3	120.5	15001	300000
mu[5]	119.2	0.1053	0.002469	118.9	119.2	119.4	15001	300000
mu[6]	118.2	0.1084	0.002599	117.9	118.2	118.4	15001	300000
mu[7]	116.4	0.1046	0.002409	116.1	116.4	116.6	15001	300000
mu[8]	114.9	0.09169	0.00178	114.6	114.9	115.0	15001	300000
mu[9]	113.5	0.07984	0.001009	113.3	113.5	113.7	15001	300000
mu[10]	112.7	0.08052	9.591E-4	112.5	112.7	112.8	15001	300000
mu[11]	111.9	0.09371	0.001679	111.7	111.9	112.1	15001	300000
mu[12]	111.1	0.1218	0.002844	110.8	111.1	111.3	15001	300000
mu[13]	110.4	0.1577	0.004114	110.0	110.4	110.7	15001	300000
sig0	0.08848	0.1029	0.001959	0.006484	0.04695	0.3729	15001	300000
siginfinity	0.07241	0.08432	0.001609	0.005299	0.03833	0.3055	15001	300000
theta	0.03813	0.002905	8.41E-5	0.03307	0.03799	0.04516	15001	300000

2. The following is a "toy version" of a large real problem that will suffice for an exam problem without being overly complicated. Consider, say, 3 shipments of some material, each consisting of 3 cartons of constant size. For

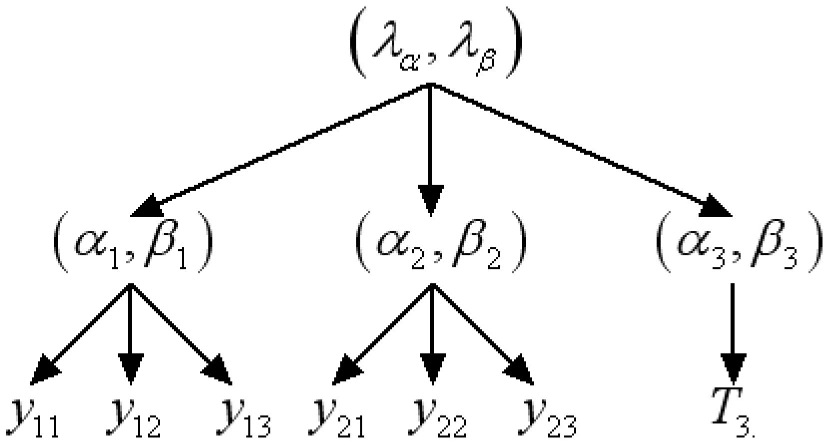
$$y_{ij} = \text{the amount of contaminant in carton } j \text{ of shipment } i$$

of interest are

$$T_{i.} = \sum_j y_{ij} \text{ for } i = 1,2,3 \text{ and } T_{..} = \sum_i T_{i.}$$

(the amount of contaminant in shipment $i = 1,2,3$ and the total amount of contaminant). We will suppose that for fixed (α_i, β_i), y_{i1}, y_{i2}, and y_{i3} are iid WinBUGS-$\Gamma(\alpha_i, \beta_i)$, and that for fixed $(\lambda_\alpha, \lambda_\beta)$, $(\alpha_1, \beta_1), (\alpha_2, \beta_2)$, and (α_3, β_3) are iid according to a joint distribution of independence with marginals that are respectively WinBUGS-Exp(λ_α) and WinBUGS-Exp(λ_β). What is observable are y_{11}, y_{12}, y_{21}, and y_{22}.

To do a Bayes analysis here, we'll suppose that *a priori* λ_α and λ_β are iid WinBUGS-Exp(1). That makes λ_α and λ_β random and makes the DAG below an appropriate representation of the joint distribution of $\left(y_{11}, y_{12}, y_{13}, y_{21}, y_{22}, y_{23}, T_{3.}, (\alpha_1, \beta_1), (\alpha_2, \beta_2), (\alpha_3, \beta_3), (\lambda_\alpha, \lambda_\beta)\right)$.



Let $f(y \mid \alpha, \beta)$ stand for the $\Gamma(\alpha, \beta)$ probability density and $g(x \mid \lambda)$ stand for the $\text{Exp}(\lambda)$ probability density.

a) Write out a joint pdf for $\left(y_{11}, y_{12}, y_{13}, y_{21}, y_{22}, y_{23}, T_{3.}, (\alpha_1, \beta_1), (\alpha_2, \beta_2), (\alpha_3, \beta_3), (\lambda_\alpha, \lambda_\beta)\right)$ using the above notation.

Suppose that $y_{11} = 1.2, y_{12} = 2.0, y_{21} = .5,$ and $y_{22} = .8$. The `WinBUGS` output on the next page is from an implementation of the above model.

```
model  {

lambdaA ~ dexp (1)
lambdaB ~ dexp (1)

for (i in 1:3) {

  alpha[i] ~ dexp (lambdaA)
  beta[i] ~ dexp (lambdaB)

  }

for (i in 1:2)  {
  for (j in 1:3)  {

    y[i,j] ~ dgamma (alpha[i],beta[i])

    }
  }

T[1] <- y[1,1]+y[1,2]+y[1,3]
T[2] <- y[2,1]+y[2,2]+y[2,3]
A3 <- 3*alpha[3]
T[3] ~ dgamma (A3,beta[3])
Tot <- T[1]+T[2]+T[3]

}

list (y = structure(.Data = c(1.2,2.0,NA,.5,.8,NA), .Dim = c(2,3))))
```

node	**mean**	**sd**	**MC error**	**2.5%**	**median**	**97.5%**	**start**	**sample**
T[1]	5.11	2.981	0.01029	3.259	4.644	9.813	25001	104607
T[2]	2.27	4.169	0.01328	1.311	1.922	5.188	25001	104607
T[3]	52.1	2131.0	6.551	5.434E-4	2.96	196.1	25001	104607
Tot	59.48	2131.0	6.55	5.591	10.59	204.0	25001	104607
lambdaA	0.5658	0.5161	0.004612	0.05638	0.4128	1.954	25001	104607
lambdaB	0.6227	0.5972	0.005276	0.05755	0.4395	2.246	25001	104607

b) What are the posterior standard deviations of $T_{1.}, T_{2.},$ and $T_{3.}$? How do they compare? Argue that their relative sizes make sense.