Stat 544 Exam 1

March 1, 2007

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

____ ________________
signature date

1. Suppose that a parameter $\boldsymbol{\theta} = (\theta_1, \theta_2)$ takes one of the values in $\left\{(1,0), \left(2, \frac{1}{4}\right), \left(2, \frac{3}{4}\right)\right\}$. Then X

is uniformly distributed on $\{0,1,2\}$ if $\theta_1 = 1$
is Binomial$(2, \theta_2)$ if $\theta_1 = 2$ (*)

A Bayesian uses prior probabilities

$$g((1,0)) = \frac{1}{2},\ g\left(\left(2, \frac{1}{4}\right)\right) = \frac{1}{4}, \text{ and } g\left(\left(2, \frac{3}{4}\right)\right) = \frac{1}{4}$$

a) In the table below give the probability mass function of the (joint) distribution of (X, θ) that the Bayesian is using. (No need to do arithmetic.)

$\boldsymbol{\theta}$	x: 0	1	2
$(1,0)$			
$\left(2, \frac{1}{4}\right)$			
$\left(2, \frac{3}{4}\right)$			

b) Give below (in tabular form: possible values and posterior probabilities) the posterior distributions of first $\boldsymbol{\theta}$ and then θ_1 if $X = 2$ (say, $g(\boldsymbol{\theta} | 2)$ and then $g(\theta_1 | 2)$).

c) Suppose that given $\boldsymbol{\theta}$, X and X_{new} are independent with marginal distributions specified as in (*). Find the posterior predictive distribution of X_{new} given that $X = 2$ and give it below in tabular form.

2. Call the p quantile of the χ^2_ν distribution $\chi^2_{\nu,p}$ and use this notation in your answers below.

a) A Bayesian observes 5 iid $\mathrm{N}(0,\sigma^2)$ observations $X_1,\ldots,X_5$ and finds that $\sum_{i=1}^{5} X_i^2 = 13$. What are 95% credible limits for σ^2 that would be reported by this person if he or she is using the Jeffreys improper prior for σ^2?

b) A Bayesian observes 5 iid $\mathrm{MVN}_2(\mathbf{0},\boldsymbol{\Sigma})$ random vectors $\mathbf{X}_1,\ldots,\mathbf{X}_5$ and finds $\mathbf{S}_0 = \sum_{i=1}^{5} \mathbf{X}_i \mathbf{X}_i' = \begin{pmatrix} 13 & 3 \\ 3 & 10 \end{pmatrix}$. What are 95% credible limits for σ_{11} the *variance* that is the upper left entry of $\boldsymbol{\Sigma}$ if he or she is using an inv-Wishart$(3,(.1\mathbf{I}))$ proper prior for $\boldsymbol{\Sigma}$?

3. The two-parameter double exponential distribution is a continuous distribution with pdf on $\Re$

$$f\left(x \mid \mu,\beta\right)=\frac{\beta}{2}\exp\left(-\beta\left|x-\mu\right|\right)$$

(this distribution is symmetric about μ, and β controls the distribution variance). Suppose that a Bayesian observes $X_1, X_2, \ldots, X_n$ that are iid according to this distribution and decides to use an improper prior for the parameter vector $\left(\mu,\beta\right)$ of the form

$$g\left(\mu,\beta\right)\propto 1\cdot\exp\left(-\gamma\beta\right)$$

for some real number γ.

a) Carefully describe a Gibbs algorithm for sampling from the posterior $g\left(\mu,\beta \mid \mathbf{x}\right)$. (Give the forms of the pdf's from which one must sample for each update. If a simulation from a standard distribution can be employed, say exactly which distribution one may use.)

b) Carefully describe any Metropolis-Hastings algorithm that can be used to sample from the posterior $g\left(\mu,\beta \mid \mathbf{x}\right)$.

4. An absent-minded gambler plays a particular slot machine daily for about 2 hours per day. This person is interested in the "win" rate for the machine (the probability, p, that any particular play results in some prize), but only records how many "wins" he sees in a day (and not exactly how may times he plays). Attached to this exam is a `WinBUGS` printout for two Bayes analyses based on

$$X_i = \text{number of wins on day } i$$

for 10 days of play ($i = 1, 2, \ldots, 10$) and consistent with the rough information that physical limitations on the machine and the gambler make it essentially certain that each

$$n_i = \text{the number of plays made on day } i$$

was somewhere between 150 and 250. (These are for Models **A** and **B**.)

a) For these two analyses, carefully (and fully) state the model assumptions used:

	Model A	Model B
(conditional) distribution of $(X_1, \ldots, X_{10}) \mid (p, n_1, \ldots, n_{10})$		
(prior) distribution of p		
(prior) distribution of $(n_1, \ldots, n_{10})$		
relation of the priors for p and $(n_1, \ldots, n_{10})$		

b) In what specific way(s) is the difference in the prior assumptions on $(n_1, \ldots, n_{10})$ reflected in the posteriors for the n_i?

c) How (if at all) is the difference in the prior assumptions on $(n_1, \ldots, n_{10})$ reflected in the posteriors for p? Why is this reasonable?

d) The gambler likes the Model A analysis. What is an interval that he believes is 95% sure to contain the number of wins in a standard 2 hour session with the machine *tomorrow*?

e) A substantially different analysis was run using what we might call Model C, with `WinBUGS` code and some summaries below. Compare these results to the Model A and Model B results and discuss why both the differences and the similarities you see are (in retrospect) reasonable.

Model C

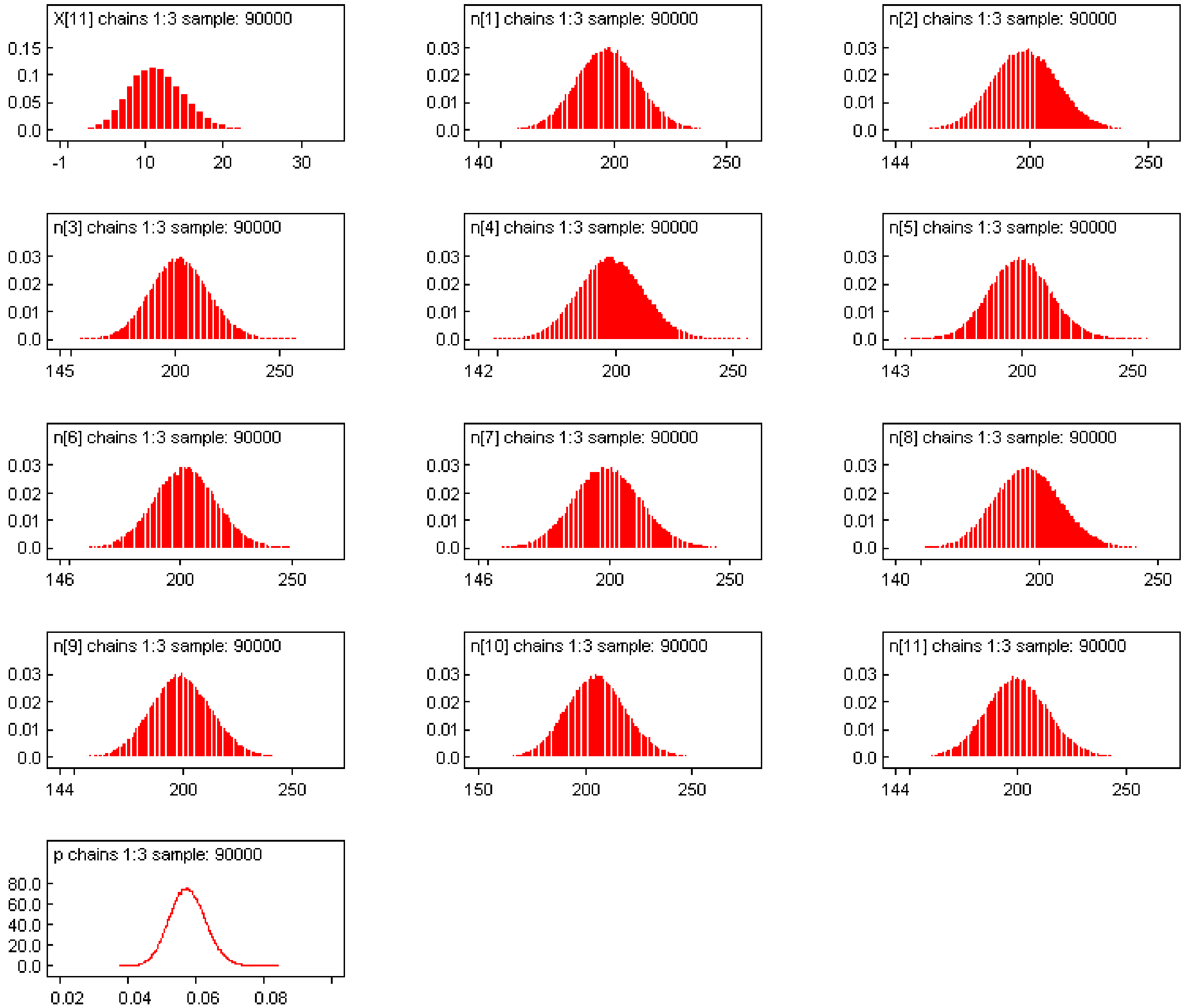
```
model {
lambda~dexp(.005)
for (i in 1:11)  {
  n[i]~dpois(lambda)
  }
p~dunif(0,1)
for (i in 1:11)  {
  X[i]~dbin(p,n[i])
  }
}
list(X=c(9,10,14,10,11,14,11,7,11,17,NA))
```

node	**mean**	**sd**	**MC error**	**2.5%**	**median**	**97.5%**	**start**	**sample**
X[11]	11.46	3.551	0.00377	5.0	11.0	19.0	61001	900000
lambda	77.66	94.39	2.194	11.91	42.12	341.7	61001	900000
n[1]	75.21	94.73	2.195	9.0	40.0	340.0	61001	900000
n[2]	76.19	94.7	2.194	10.0	41.0	341.0	61001	900000
⋮								
n[10]	83.19	94.7	2.194	17.0	48.0	348.0	61001	900000
n[11]	77.66	94.79	2.194	10.0	42.0	343.0	61001	900000
p	0.3493	0.2683	0.005312	0.03306	0.2715	0.942	61001	900000

Model A

```
model {

for (i in 1:11)  {
  n[i]~dpois(200)
  }

p~dbeta(1,1)

for (i in 1:11)  {
  X[i]~dbin(p,n[i])
  }

}

list(X=c(9,10,14,10,11,14,11,7,11,17,NA))
```

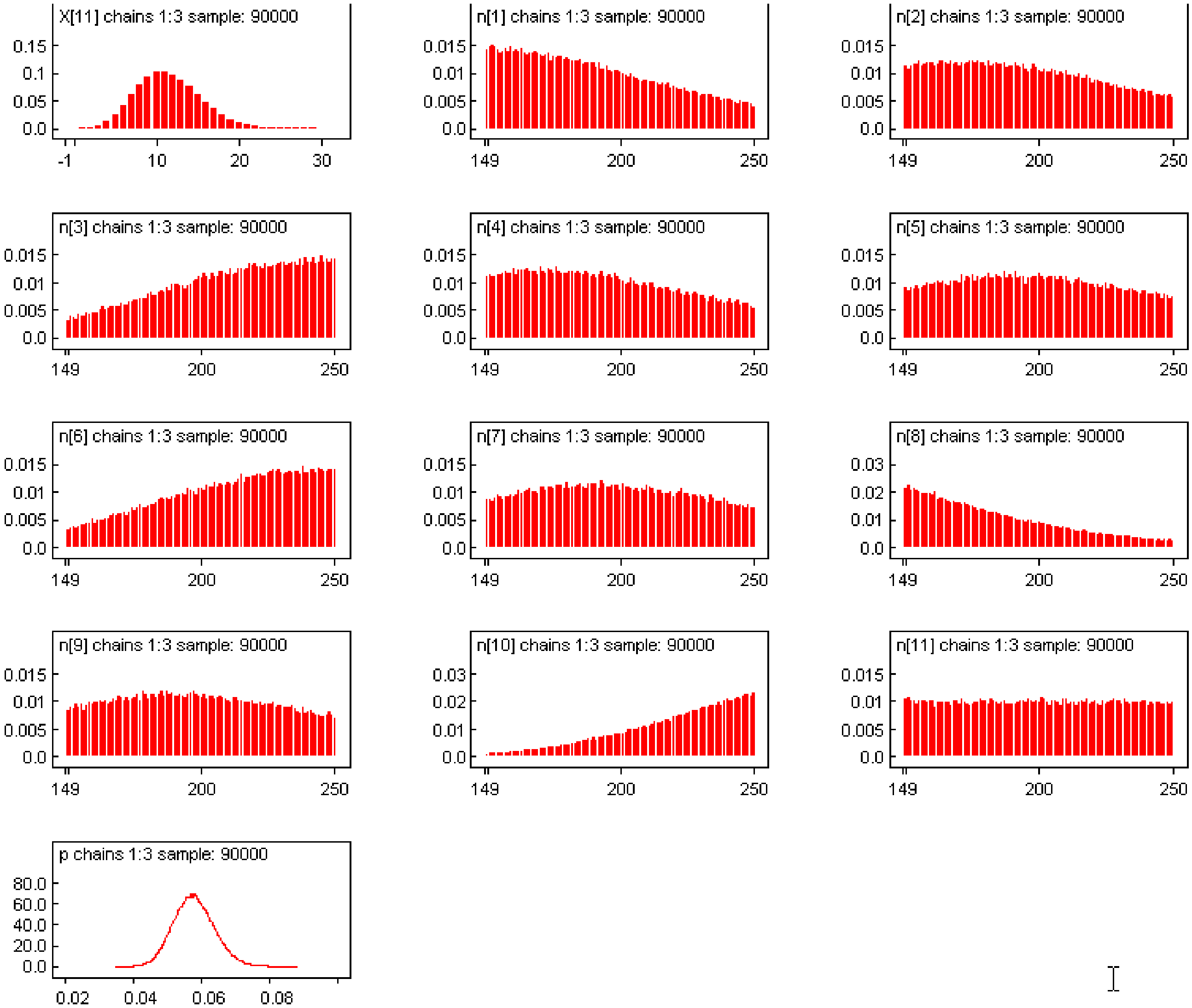
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
X[11]	11.51	3.545	0.01181	5.0	11.0	19.0	31001	90000
n[1]	197.4	13.78	0.04966	171.0	197.0	225.0	31001	90000
n[2]	198.5	13.8	0.04444	172.0	198.0	226.0	31001	90000
n[3]	202.5	13.75	0.04545	176.0	202.0	230.0	31001	90000
n[4]	198.5	13.7	0.04662	172.0	198.0	226.0	31001	90000
n[5]	199.4	13.8	0.04601	173.0	199.0	227.0	31001	90000
n[6]	202.5	13.83	0.04462	176.0	202.0	230.0	31001	90000
n[7]	199.5	13.77	0.04515	173.0	199.0	227.0	31001	90000
n[8]	195.5	13.74	0.04557	169.0	195.0	223.0	31001	90000
n[9]	199.5	13.76	0.04263	173.0	199.0	227.0	31001	90000
n[10]	205.5	13.75	0.04313	179.0	205.0	233.0	31001	90000
n[11]	199.9	14.11	0.04914	173.0	200.0	228.0	31001	90000
p	0.05755	0.00535	1.953E-5	0.04747	0.0574	0.06847	31001	90000


X[11] chains 1:3 sample: 90000
0.15
0.1
0.05
0.0
-1
10
20
30
n[1] chains 1:3 sample: 90000
0.03
0.02
0.01
0.0
140
200
250
n[2] chains 1:3 sample: 90000
144
200
250
n[3] chains 1:3 sample: 90000
145
200
250
n[4] chains 1:3 sample: 90000
142
200
250
n[5] chains 1:3 sample: 90000
143
200
250
n[6] chains 1:3 sample: 90000
146
200
250
n[7] chains 1:3 sample: 90000
146
200
250
n[8] chains 1:3 sample: 90000
140
200
250
n[9] chains 1:3 sample: 90000
144
200
250
n[10] chains 1:3 sample: 90000
150
200
250
n[11] chains 1:3 sample: 90000
144
200
250
p chains 1:3 sample: 90000
80.0
60.0
40.0
20.0
0.0
0.02
0.04
0.06
0.08

Model B

```
model {

for (j in 1:101) {
  alpha[j]<-(1/101)
  }

for (i in 1:11)  {
  m[i]~dcat(alpha[])
  n[i]<-m[i]+149
  }

p~dunif(0,1)

for (i in 1:11)  {
  X[i]~dbin(p,n[i])
  }

}

list(X=c(9,10,14,10,11,14,11,7,11,17,NA))
```

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
X[11]	11.54	3.888	0.01362	5.0	11.0	20.0	31001	90000
n[1]	190.3	27.21	0.09146	151.0	187.0	245.0	31001	90000
n[2]	194.4	27.66	0.09497	152.0	192.0	246.0	31001	90000
n[3]	210.0	26.54	0.0897	156.0	213.0	249.0	31001	90000
n[4]	194.2	27.67	0.09264	152.0	192.0	246.0	31001	90000
n[5]	198.2	27.87	0.09913	152.0	197.0	247.0	31001	90000
n[6]	210.0	26.54	0.09091	156.0	214.0	249.0	31001	90000
n[7]	198.4	27.79	0.09596	152.0	198.0	247.0	31001	90000
n[8]	183.1	25.41	0.08395	151.0	177.0	241.0	31001	90000
n[9]	198.3	27.81	0.09577	152.0	197.0	247.0	31001	90000
n[10]	219.6	23.33	0.08045	165.0	225.0	249.0	31001	90000
n[11]	199.9	29.15	0.09933	152.0	200.0	248.0	31001	90000
p	0.05768	0.005921	2.393E-5	0.04678	0.05747	0.06995	31001	90000


X[11] chains 1:3 sample: 90000
0.15
0.1
0.05
0.0
-1
10
20
30
n[1] chains 1:3 sample: 90000
n[2] chains 1:3 sample: 90000
n[3] chains 1:3 sample: 90000
n[4] chains 1:3 sample: 90000
n[5] chains 1:3 sample: 90000
n[6] chains 1:3 sample: 90000
n[7] chains 1:3 sample: 90000
n[8] chains 1:3 sample: 90000
n[9] chains 1:3 sample: 90000
n[10] chains 1:3 sample: 90000
n[11] chains 1:3 sample: 90000
0.015
0.01
0.005
0.03
0.02
149
200
250
p chains 1:3 sample: 90000
80.0
60.0
40.0
20.0
0.0
0.02
0.04
0.06
0.08