Stat 544 Exam 1

March 13, 2008

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

____
signature

____
name printed

There are 12 parts on this exam. I will score every part out of 10 points and take your best 10 of 12 scores. (Budget your time accordingly.)

1. Below are possible distributions for Y depending upon the value of a parameter $\theta \in \{1,2,3\}$.

y	$f(y\mid 1)$		y	$f(y\mid 2)$		y	$f(y\mid 3)$
1	$1/2$		1	$1/4$		1	0
2	$1/2$		2	$1/2$		2	$1/2$
3	0		3	$1/4$		3	$1/2$

a) Two independent observations Y_1 and Y_2 (with marginal pmfs $f(y\mid\theta)$) take values 2 and 3. Based on this an a prior for θ that is uniform on $\{1,2,3\}$, find the posterior distribution for θ.

b) Based on your answer to a) find a posterior predictive distribution for Y_{new} (that conditional on θ is independent of Y_1 and Y_2 with the same marginal pmf).

c) Bayesians like to argue that as long as their priors "spread the prior probability around sufficiently" (so that they are sure to "cover" any true parameter value) their posteriors will be "consistent" in the sense of piling up around the true value of a parameter with increasing sample information. Suppose that in fact $\theta = 2$ and $Y_1, Y_2, \ldots$ are iid with marginal pmf $f(\cdot \mid 2)$. For the uniform prior on $\{1,2,3\}$ here evaluate

$$P_{\theta=2}\left[g(1 \mid Y_1,\ldots,Y_n) > 0\right] \text{ and } P_{\theta=2}\left[g(3 \mid Y_1,\ldots,Y_n) > 0\right]$$

(These are $\theta = 2$ probabilities that through n observations, the posterior has not eliminated respectively the possibilities that $\theta = 1$ and $\theta = 3$.)

d) (Censoring) Suppose that in contrast to the situation in part a), I get to see not the full information in (Y_1, Y_2), but only the values of

$$Z_i = \begin{cases} 0 & \text{if } Y_i = 1 \\ 1 & \text{if } Y_i = 2 \text{ or } 3 \end{cases}$$

for $i = 1,2$. Based still on the uniform prior over $\{1,2,3\}$, but now the information that $Z_1 = 1$ and $Z_2 = 1$, what is the posterior distribution of θ?

2. Consider the possibility of specifying a "distribution" for (θ_1, θ_2) by a "density" on $\Re^2$ by

$$g(\theta_1, \theta_2) \propto \exp\left(-(\theta_1 - \theta_2)^2\right)$$

a) Argue very carefully that such a $g(\cdot)$ does NOT specify a proper probability distribution for (θ_1, θ_2).

b) In spite of the fact in a), someone who wasn't paying attention might attempt to "sample from $g(\cdot)$" using a Gibbs sampler. Identify exactly how that person would make updates (from, say, (θ_1^i, θ_2^i) to $(\theta_1^{i+1}, \theta_2^{i+1})$). (What are the two conditionals from which the person would sample?)

c) How do you expect the sampler described in b) to behave (from, say, a start at $(0,0)$)? Do you expect the (joint) relative frequency distribution of $\left\{\left(\theta_1^i,\theta_2^i\right)\right\}_{i=1,\ldots,N}$ to converge? If so, what can you say about the limit? If not, what can you say about what you expect to happen to the relative frequency distribution with increasing N?

d) In spite of the fact in a), in a model where conditioned on (θ_1,θ_2) variables Y_1 and Y_2 are independent normal variables with means θ_1 and θ_2 and variance 1, $g(\cdot)$ can be used as an improper prior and produces a legitimate posterior distribution. Carefully argue that this "works." (Hint: You can bound $g(\cdot)$ above by some constant.)

3. Suppose that conditioned on positive parameters $\lambda, \delta_1,$ and δ_2, the variables $Y_1, Y_2,$ and Y_3 are independent Poisson variables with respective means $\lambda, \lambda + \delta_1,$ and $\lambda + \delta_1 + \delta_2$. Suppose that one observes $Y_1 = 7, Y_2 = 5,$ and $Y_3 = 6$ and uses independent $\mathrm{U}(0,10)$ priors for $\lambda, \delta_1,$ and δ_2.

a) Describe completely a Metropolis-Hastings or a Metropolis-Hastings-within-Gibbs algorithm that one could use to sample from the posterior distribution of the parameters.

b) Write `WinBUGS` code for implementing a Gibbs sampler for the posterior distribution of the parameters.

4. Suppose that someone is interested in the mean of the determinant of a sample covariance matrix for $n = 20$ observations from $\text{MVN}_3(\boldsymbol{\mu}, 2\mathbf{I})$. Write `WinBUGS` code that could be used to evaluate this mean. (Degrees of freedom for the sample covariance matrix are 19 and `WinBUGS` has a log-determinant function `logdet()`.)

5. Some angles between holes drilled (using so-called electrical discharge machining) in precision metal parts of a certain type and the flat top surfaces of the parts are supposed to be $45^\circ \pm 2^\circ$. 10 such measured angles are below (measurements to the nearest degree)

$$46, 45, 45, 45, 44, 45, 43, 45, 45, 46$$

Supposing that angles are iid $\mathrm{N}\left(\mu, \sigma^2\right)$, some quantities potentially of interest to the manufacturer are

$$Y_{\text{new}}$$

$$p\left(\mu, \sigma^2\right) = \Phi\left(\frac{47-\mu}{\sigma}\right) - \Phi\left(\frac{43-\mu}{\sigma}\right)$$

$$C_p = \frac{47-43}{6\sigma}$$

$$C_{pk} = \min\left\{\frac{47-\mu}{3\sigma}, \frac{\mu-43}{3\sigma}\right\} = \frac{(47-43)-2\left|\mu-45\right|}{6\sigma}$$

(an additional angle, the fraction of angles that conform to the requirements, a ratio of spread in requirements to the "spread" in the distribution of angles, and a so-called "process capability ratio").

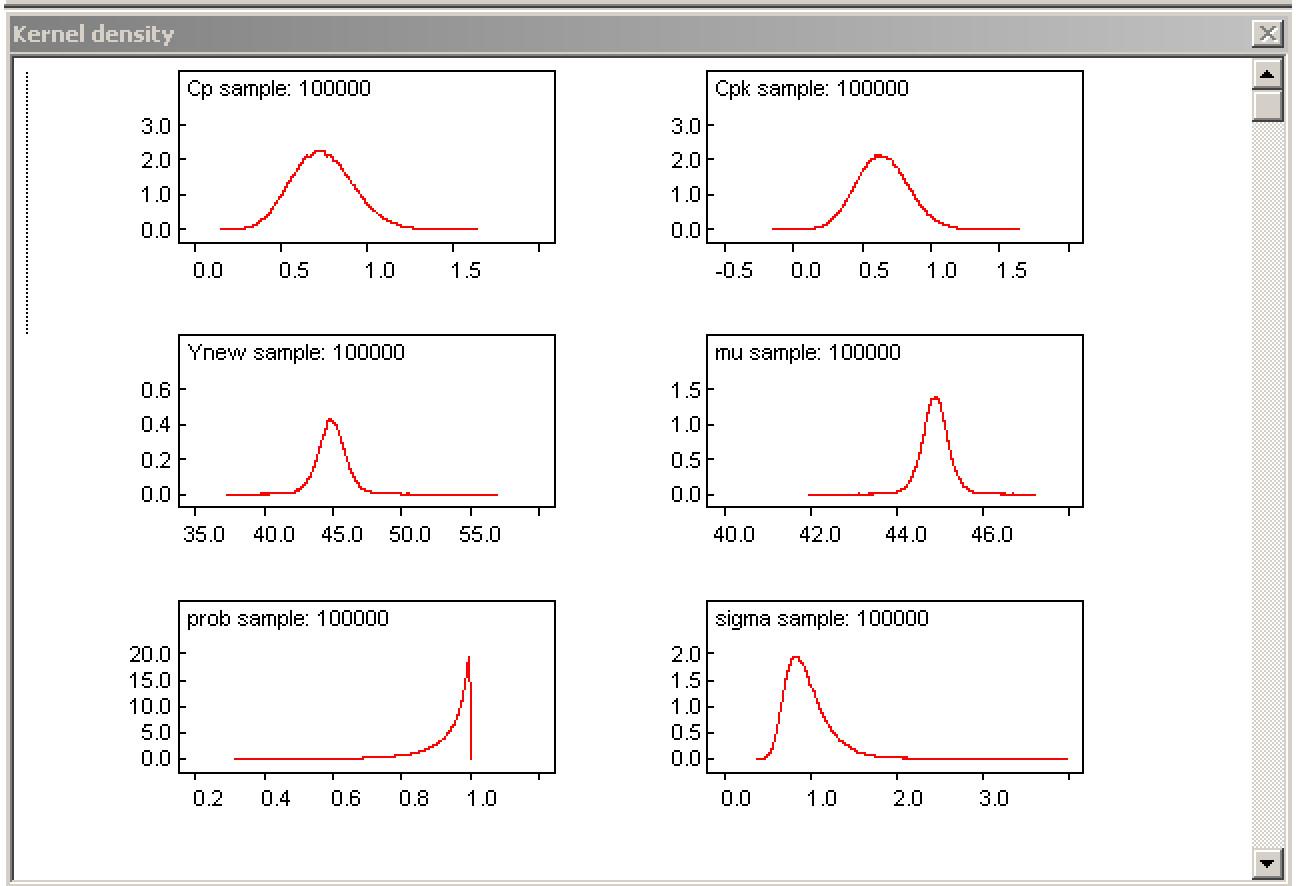
There are two `WinBUGS` printouts for analyses of the observed angles following this page. What do these indicate about the extent to which the manufacturer is producing parts with angles meeting the $45^\circ \pm 2^\circ$ requirements?

```
model {
mu~dflat()
logsigma~dflat()
sigma<-exp(logsigma)
tau<-exp(-2*logsigma)
for (i in 1:10) {
Y[i]~dnorm(mu,tau)
  }
Ynew~dnorm(mu,tau)
prob<-phi((47-mu)/sigma)-phi((43-mu)/sigma)
Cp<-2/(3*sigma)
Cpk<-((4-2*abs(mu-45))/(6*sigma))
}

list(Y=c(46,45,45,45,44,45,43,45,45,46))
list(mu=45,logsigma=0,Ynew=45)
```

Node statistics

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
Cp	0.7404	0.1767	5.864E-4	0.4183	0.7331	1.106	111001	100000
Cpk	0.651	0.1868	6.546E-4	0.3017	0.6455	1.031	111001	100000
Ynew	44.9	1.043	0.003294	42.83	44.9	46.98	111001	100000
mu	44.9	0.3134	9.615E-4	44.27	44.9	45.53	111001	100000
prob	0.9417	0.0649	2.273E-4	0.7611	0.9643	0.9987	111001	100000
sigma	0.9582	0.2597	8.77E-4	0.603	0.9093	1.594	111001	100000



```
model {
mu~dflat()
logsigma~dflat()
sigma<-exp(logsigma)
tau<-exp(-2*logsigma)
X[1]~dnorm(mu,tau) I(45.5,46.5)
X[2]~dnorm(mu,tau) I(44.5,45.5)
X[3]~dnorm(mu,tau) I(44.5,45.5)
X[4]~dnorm(mu,tau) I(44.5,45.5)
X[5]~dnorm(mu,tau) I(43.5,44.5)
X[6]~dnorm(mu,tau) I(44.5,45.5)
X[7]~dnorm(mu,tau) I(42.5,43.5)
X[8]~dnorm(mu,tau) I(44.5,45.5)
X[9]~dnorm(mu,tau) I(44.5,45.5)
X[10]~dnorm(mu,tau) I(45.5,46.5)
Xnew~dnorm(mu,tau)
Ynew<-round(Xnew)
prob<-phi((47-mu)/sigma)-phi((43-mu)/sigma)
Cp<-2/(3*sigma)
Cpk<-((4-2*abs(mu-45))/(6*sigma))
}

list(mu=45,logsigma=0,Xnew=45)
```

Node statistics

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
Cp	0.7955	0.2109	8.886E-4	0.4281	0.7798	1.25	111001	100000
Cpk	0.6994	0.2169	9.414E-4	0.31	0.686	1.159	111001	100000
Ynew	44.89	1.035	0.003177	43.0	45.0	47.0	111001	100000
mu	44.89	0.3118	0.001081	44.27	44.89	45.52	111001	100000
prob	0.9503	0.06309	2.656E-4	0.7702	0.9739	0.9997	111001	100000
sigma	0.9035	0.2681	0.00111	0.5333	0.8549	1.557	111001	100000