ECON 546: Themes in Econometrics

Examples of Markov Chain Monte Carlo Analysis

Example 1

Let X and Y be a pair of random variables whose joint distribution is

$$f(x, y) \propto ({}^{n}C_{x}) y^{x+\alpha-1}(1-y)^{n-x+\beta-1} \quad ; \quad x = 0, 1, 2, 3, \ldots, \text{n} \; ; \; 0 < y < \infty$$

(a) What is the conditional distribution for X, given Y?
(b) What is the conditional distribution for Y, given X?
(c) Let's use the EViews workfile **S:\Social Sciences\Economics\ECON546\MCMC1.wf1** and the program file **S:\Social Sciences\Economics\ECON546\MCMC1.prg** to obtain the *marginal* distributions of X and Y. Establish that $E(X)$ = 3.3, $E(Y)$ = 0.33, S.D.(X) = 2.246, and S.D. (Y) = 0.178. These results do *not* depend on the initial value for Y.

```
rndseed 123456
!nrep=50000
vector (!nrep) margy
vector (!nrep) margx
scalar n=10
scalar y=0.1
scalar a=2
scalar b=4
for !i=1 to !nrep
  scalar x=@rbinom(n,y)
  margx(!i)=x
  scalar y=@rbeta(x+a,n-x+b)
  margy(!i)=y
next
smpl 1 !nrep
mtos(margx,px)
mtos(margy,py)
smpl 1001 !nrep
px.hist
py.hist
```

Example 2

We'll use the EViews workfile, **S:\Social Sciences\Economics\ECON546\MCMC2.wf1** and the EViews program file **S:\Social Sciences\Economics\ECON546\MCMC2.prg** for this example.

Here is the estimation problem:

We have a sample of 10 "count data" observations that have been generated by *two* Poisson processes. The first m values come from a Poisson distribution with *unknown mean*, θ_1, and the rest of the observations come from an independent Poisson distribution with *known mean* of θ_2. The interesting part of the problem is that the value of m is also *unknown*.

As θ_1 must be positive, we put a Gamma prior on this parameter. We put a uniform prior on m over the range 1 to 10:

$$p(\theta_1) \propto (\theta_1 / b_1)^{r_1 - 1} e^{-\theta_1 / b_1}$$

$$p(m) = 1/n$$

The likelihood function is:

$$L = \prod_{i=1}^{m} [\theta_1^{y_i} e^{-\theta_1} / y_i!] \prod_{i=m+1}^{n} [\theta_2^{y_i} e^{-\theta_2} / y_i!] \propto e^{-m\theta_1 - (n-m)\theta_2} [\theta_1^{\sum_1^m y_i}][\theta_2^{\sum_{m+1}^n y_i}] .$$

So, recalling that θ_2 is known, the joint posterior for the parameters is:

$$p(\theta_1, m \mid \theta_2, y) \propto (1/n)(\theta_1 / b_1)^{r_1 - 1} e^{-\theta_1 / b_1} e^{-m\theta_1 - (n-m)\theta_2} [\theta_1^{\sum_1^m y_i}][\theta_2^{\sum_{m+1}^n y_i}] .$$

The *conditional* posterior density for θ_1 is:

$$p(\theta_1 \mid m, \theta_2, y) \propto \theta_1^{r_1 - 1 + \sum_1^m y_i} e^{-\theta_1 (m + 1/b_1)} ,$$

which is Gamma, with parameters [1/(m + (1/b_1))] and $[r_1 + \sum_{i=1}^{m} y_i]$.

The *conditional* posterior for m is a completely *non-standard* p.m.f. on [1 , n]:

$$p(m \mid \theta_1, \theta_2, y) \propto \exp[m(\theta_1 - \theta_2)](\theta_1 / \theta_2)^{\sum_1^m y_i} .$$

We can simulate drawings from this discrete distribution by using the so-called "Table Lookup Method".

The program file, **MCMC2.prg** (see the next two pages of this handout) applies the Gibbs sampler to determine the marginal posteriors for m and θ_1, which enables us to obtain Bayes estimators for these two parameters.

```
' PROGRAM TO UNDERTAKE BAYESIAN ANALYSIS OF A POISSON PROCESS WITH A
CHANGE-POINT AT AN UNKNOWN LOCATION IN THE DATA
' THE COUNT DATA RANGE FROM 0 TO 10 IN VALUE

' THE SECOND PART OF THE PROCESS IS KNOWN TO BE POISSON WITH A MEAN OF 6

' A GAMMA PRIOR IS USED FOR THE FIRST POISSON MEAN, AND A UNIFORM PRIOR IS
USED FOR THE CHANGE-POINT

!nrep=1000
' INITIALIZE SOME VALUES
'********************************
rndseed 123456
' SET THE PRIOR MEAN FOR THETA1 TO BE 6, AND THE PRIOR VARIANCE TO BE 0.06
(SD = 0.245)
scalar r1=600
scalar b1=.01
scalar t2=6

vector(!nrep) margt1
vector(!nrep) margm
vector(10) pm
vector(10) cusum
vector(10) mm
smpl 1 10
scalar t1=1

scalar m
smpl 1 10
series sumofy=@sum(y)
series sumy=0
series sumy1

'******************************
'START OF THE MCMC LOOP
for !i=1 to !nrep

' GENERATE A NEW VALUE FOR M USING THE "TABLE LOOKUP METHOD"
' FIRST CONSTRUCT THE CONDITIONAL POSTERIOR P.D.F. AND C.D.F. FOR M

scalar sum=0
scalar sump=0
for !j=1 to 10
smpl 1 1+!j-1
series sigy=@sum(y)
scalar siggy=sigy(1)
mm(!j)=exp(!j*(t1-t2))*((t1/t2)^siggy)
sum=sum+mm(!j)
next
for !j = 1 to 10
pm(!j)=mm(!j)/sum
sump=sump+pm(!j)
cusum(!j)=sump
next
```

```
' NOW IMPLEMENT THE "TABLE LOOKUP METHOD" ITSELF TO ACTUALLY GENERATE
RANDOM M VALUES FROM THE NON-STANDARD CONDITIONAL POSTERIOR
DISTRIBUTION

scalar u=@runif(0,1)
for !j = 1 to 10
m=!j
if  u < cusum(!j) then
!j=10
endif
next
margm(!i)=m


' NOW GENERATE THE CONDITIONAL POSTERIOR FOR THETA1

sumy=0
for !k=1 to m
sumy=sumy+@elem(y,!k)
next
sumy1=sumofy-sumy
scalar par1=1/(m+1/b1)
scalar par2=r1+sumy(1)
 t1=@rgamma(par1, par2)
margt1(!i)=t1

next

' END OF THE MCMC LOOP

smpl 1 !nrep
' CONVERT VECTORS TO SERIES TO FACILITATE PLOTS, ETC.
mtos(margm,postm)
mtos(margt1,postt1)
mtos(pm,pms)

' ALLOW FOR "BURN-IN" PERIOD
‘ IDEALLY, !NREP SHOULD BE MUCH LARGER & WE WOULD HAVE A LARGER “BURN-IN”
smpl 201 !nrep
postm.hist
postt1.hist
```