MATH 598: TOPICS IN STATISTICS

SIMULATING MARKOV CHAINS

Simulation 1: Here

$$P = \begin{pmatrix} 0.3 & 0.7 \\ 0.9 & 0.1 \end{pmatrix}$$

For detailed balance, we require that

$$\pi_1 P_{12} = (1-\pi_1)P_{21} \qquad \text{or} \qquad \frac{\pi_1}{1-\pi_1} = \frac{P_{21}}{P_{12}} = \frac{0.9}{0.7}$$

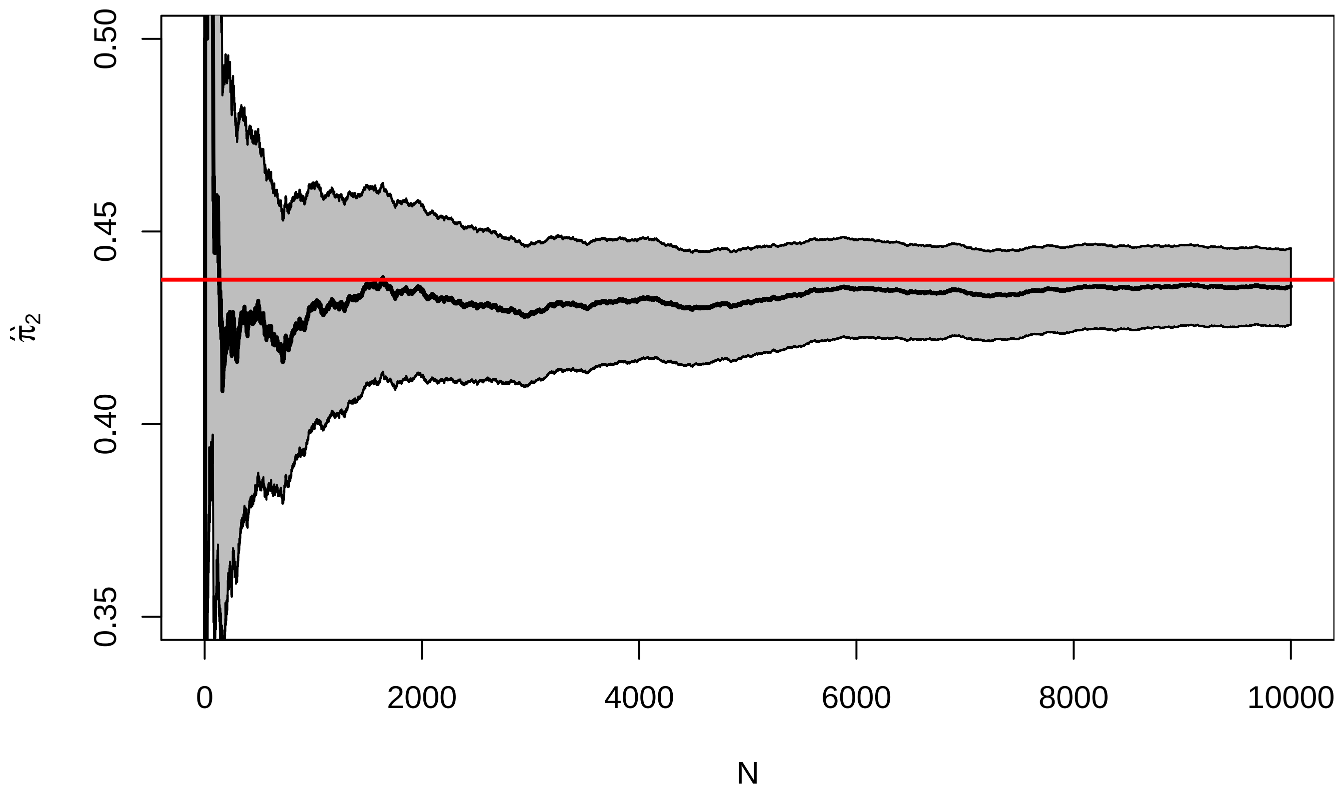
so therefore have that

$$\pi_1 = \frac{0.9}{0.7+0.9} = 0.562500 \qquad \pi_2 = 1 - \pi_1 = 0.437500.$$

This chain is therefore reversible with respect to this π, and we have that

$$\mathbf{1}\pi = \lim_{k\longrightarrow\infty} P^k.$$

```
set.seed(3242)
d<-2
P<-matrix(c(0.3,0.7,0.9,0.1),d,d,byrow=T)
P

+      [,1] [,2]
+ [1,]  0.3  0.7
+ [2,]  0.9  0.1

N<-10000
X<-rep(0,N)
X[1]<-1
for(i in 2:N){
        X[i]<-sample(c(1:d),size=1,prob=P[X[i-1],])
}

Ptmp<-P
for(k in 1:200){Ptmp<-P%*% Ptmp  }
Ptmp

+        [,1]   [,2]
+ [1,] 0.5625 0.4375
+ [2,] 0.5625 0.4375

Pi<-Ptmp[1,]
P.est<-cumsum(X-1)/c(1:N)
P.var<-P.est*(1-P.est)/c(1:N)
par(mar=c(4,4,2,0))
plot(c(1:N),xlab='N',P.est,ylim=range(0.35,0.5),type="n",ylab=expression(hat(pi)[2]))
polygon(c(c(1:N),rev(c(1:N))),c(P.est+2*sqrt(P.var),rev(P.est-2*sqrt(P.var))),col="gray")
lines(c(1:N),P.est,lwd=2)
abline(h=Pi[2],col="red",lwd=2)
```



Simulation 2: Now suppose we wish to design a chain for a two-state problem such that it has a stationary distribution $(\pi_1, \pi_2) = (0.2, 0.8)$. From the construction in lectures, we may specify

$$P_{12} = \min\left\{1, \frac{1-\pi_1}{\pi_1}\right\} \qquad P_{21} = \min\left\{1, \frac{\pi_1}{1-\pi_1}\right\}$$

that is

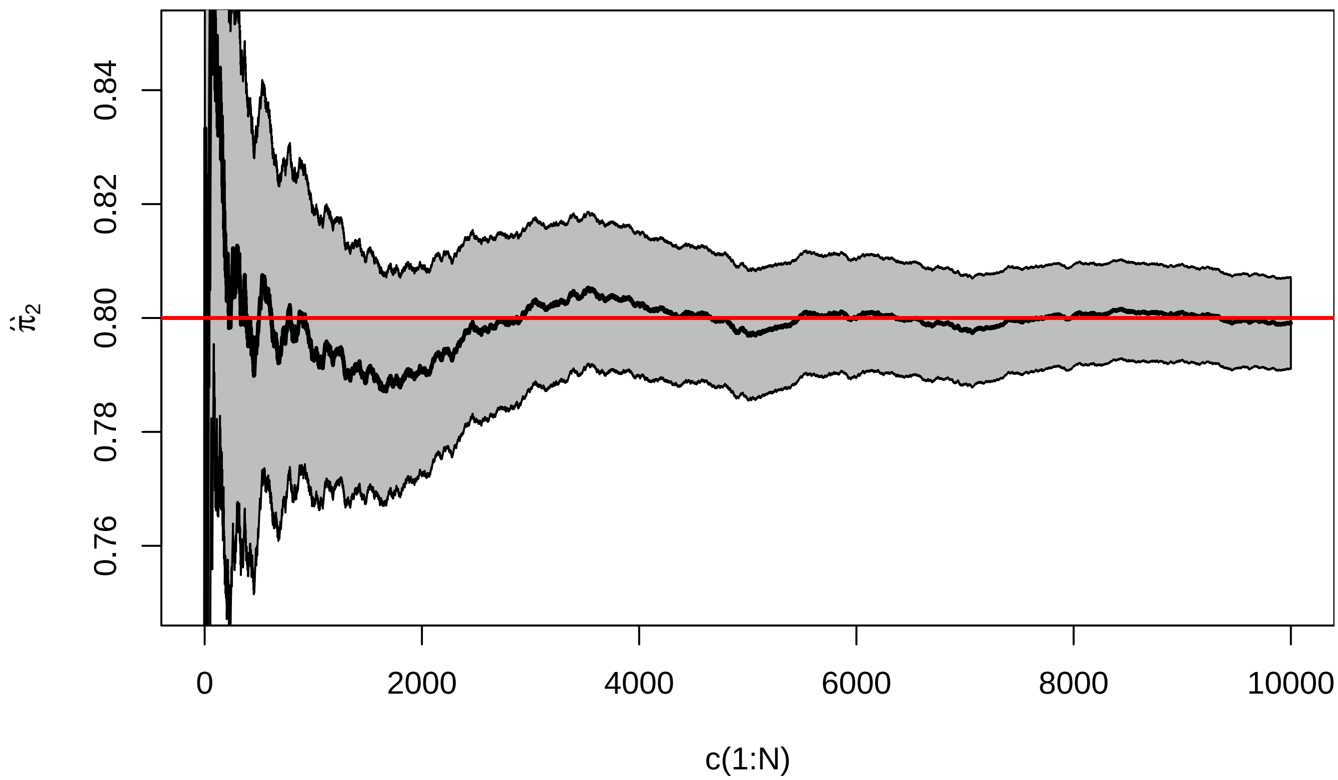
$$P_{12} = \min\left\{1, \frac{0.8}{0.2}\right\} = 1 \qquad P_{21} = \min\left\{1, \frac{0.2}{0.8}\right\} = 0.25$$

and therefore need

$$P = \begin{pmatrix} 0.00 & 1.00 \\ 0.25 & 0.75 \end{pmatrix}$$

```
set.seed(242)
d<-2
P<-matrix(c(0,1,0.25,0.75),d,d,byrow=T)
P

+      [,1] [,2]
+ [1,] 0.00 1.00
+ [2,] 0.25 0.75

N<-10000
X<-rep(0,N)
X[1]<-1
for(i in 2:N){
        X[i]<-sample(c(1:d),size=1,prob=P[X[i-1],])
}

Ptmp<-P
for(k in 1:200){Ptmp<-P%*% Ptmp  }
Ptmp
```

```
+       [,1] [,2]
+ [1,]  0.2  0.8
+ [2,]  0.2  0.8

Pi<-Ptmp[1,]
P.est<-cumsum(X-1)/c(1:N)
P.var<-P.est*(1-P.est)/c(1:N)
par(mar=c(4,4,1,0))
plot(c(1:N),P.est,ylim=range(0.75,0.85),type="n",ylab=expression(hat(pi)[2]))
polygon(c(c(1:N),rev(c(1:N))),c(P.est+2*sqrt(P.var),rev(P.est-2*sqrt(P.var))),col="gray")
lines(c(1:N),P.est,lwd=2)
abline(h=Pi[2],col="red",lwd=2)
```



Simulation 3: Poisson distribution simulation using Metropolis-Hastings algorithm.
Suppose we wish to simulate from the discrete distribution $(\pi_0, \pi_1, \pi_2, \ldots)$, where for $\lambda > 0$,

$$\pi_i = \frac{e^{-\lambda}\lambda^i}{i!} \qquad i = 0, 1, 2, \ldots.$$

Following the Metropolis-Hastings (MH) formulation, let matrix Q define the proposal probabilities

$$[Q]_{ij} = \Pr[Z_t = j | X_t = i]$$

and define the *acceptance probabilities*

$$\alpha_{ij} = \min\left\{1, \frac{\pi_j}{\pi_i}\frac{q_{ji}}{q_{ij}}\right\}$$

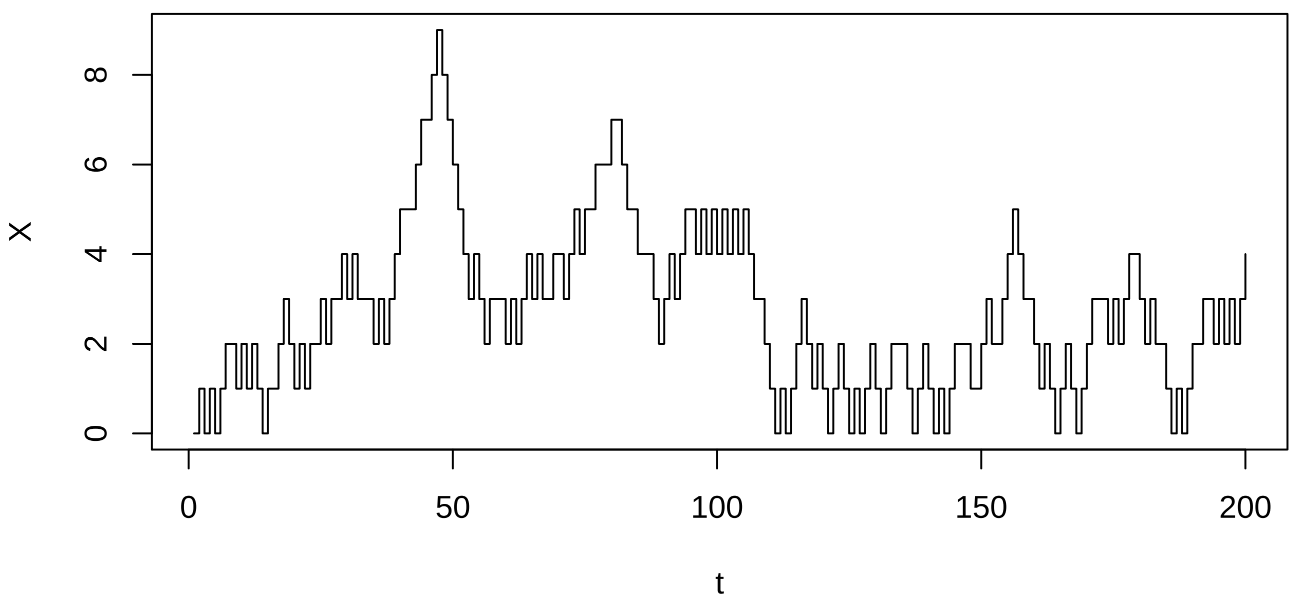
In the MH Markov chain, if $X_t = i$, we set $X_{t+1} = Z_t = j$ with probability α_{ij}. Otherwise we set $X_{t+1} = X_t = i$. Suppose

$$q_{ij} = \begin{cases} 1 & i = 0, j = 1 \\ \frac{1}{2} & i \geq 1, j = i-1, i+1 \\ 0 & \text{otherwise} \end{cases}$$

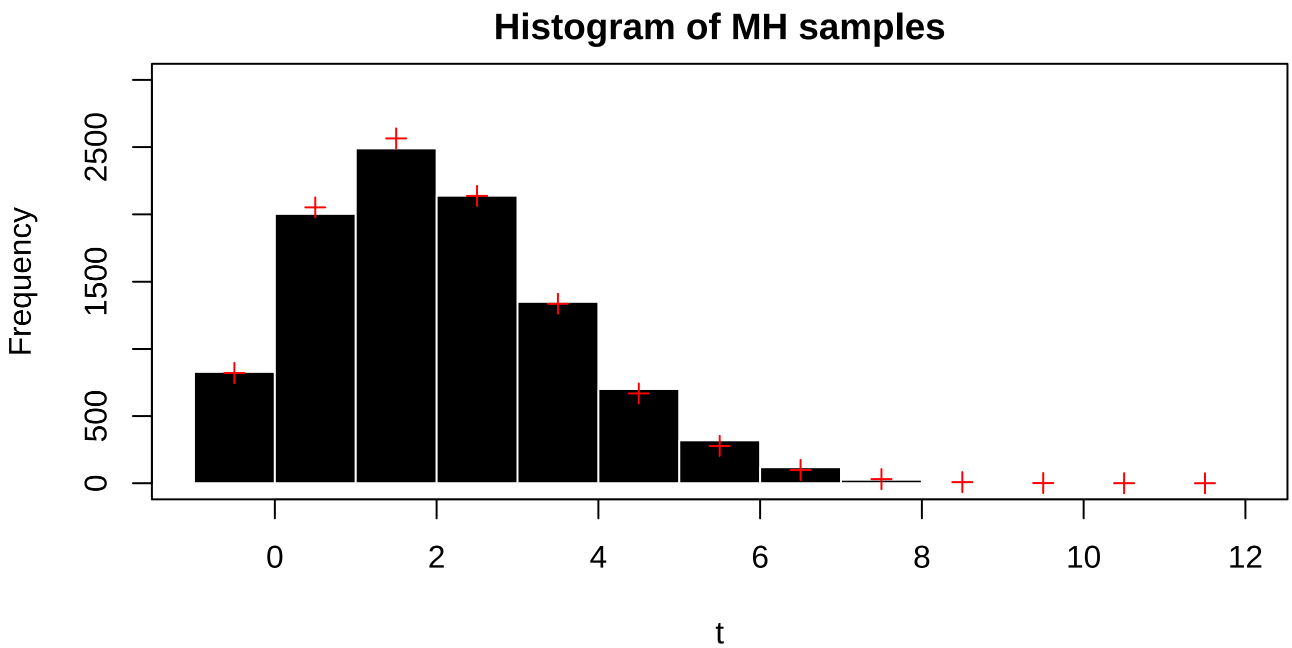
That is, Z_t is proposed uniformly on the finite set $\{x_t - 1, x_t + 1\}$, unless $X_t = 0$, in which case $Z_t = 1$ is proposed.

The simulation proceeds as follows:

```
set.seed(4532)
N<-10000
lam<-2.5
X<-rep(0,N)
X[1]<-0
for(istep in 2:N){
        if(X[istep-1] == 0){
                Z<-1
                qrat<-1/2
                al<-min(1,qrat*dpois(Z,lam)/dpois(X[istep-1],lam))
        }else{
                Z<-sample(c(X[istep-1]-1,X[istep-1]+1),size=1,prob=c(1/2,1/2))
                if(Z == 0){
                        qrat<-2
                }else{
                        qrat<-1
                }
                al<-min(1,qrat*dpois(Z,lam)/dpois(X[istep-1],lam))
        }
        u<-runif(1)

        if(u < al){
                X[istep]<-Z
        }else{
                X[istep]<-X[istep-1]
        }
}
par(mar=c(4,4,2,1))
plot(X[1:200],type="s",xlab="t",ylab="X")
```




```
hist(X,br=c(-1:12),xlab="t",main="Histogram of MH samples",ylim=range(0,3000),col="black",border="white");b
points(c(0:12)-0.5,dpois(c(0:12),lam)*N,col="red",pch=3)
```




```
iburn<-1000
tvn<-c(iburn:N)*0
Xlim<-100
for(i in iburn:N){
        Xtab<-table(X[1:i])
        pihat<-Xtab/i
        pivals<-as.numeric(names(Xtab))
        pihatvec<-rep(0,Xlim);pihatvec[pivals+1]<-pihat
        pivec<-dpois(c(0:(Xlim-1)),lam)
        tvn[i-iburn+1]<-sum(abs(pivec-pihatvec))/2
}
plot(c(iburn:N),tvn,ylim=range(0,0.1),type="l",xlab="N",ylab="TV")
title('Total variation distance from truth')
```