Statistical Decision Theory

Introduction

This tutorial provides R code for using statistical decision theory (SDT) to decide on a management action for the Henslow's sparrow example presented in the Bayesian model selection and decision theory workshop held at the international Statistical Ecology Conference on Sunday June 26, 2016. The example first appeared in Williams and Hooten (2016). Henslow's sparrow data can be found here. We assume that the reader is familiar with implementing Bayesian generalized linear models in R or Jags. A custom MCMC sampler is used to fit a model. However, equivalent Jags code could be used instead.

Our example involves a refuge manager deciding on a prescribed-burn interval to implement in a management plan for Henslow's sparrows on eight grasslands over the next 20 years. The problem was motivated by a real management issue at Big Oaks National Wildlife Refuge in southeastern Indiana, USA from which our data arise. The manager has information on: the choice of actions for a burn interval $\delta \in \{1, 2, 3, 4\}$, an average annual abundance of Henslow's sparrows $N_{\delta,\boldsymbol{\theta}}$ that depends on unknown model parameters $\boldsymbol{\theta}$ and the choice of management action δ, data on the abundance of Henslow's sparrows (y) at eight sites for a four-year period following prescribed fire, prior information from another study on the correlation between Henslow's sparrow abundance and prescribed fire, and a perception of loss that includes the expense of the management action, and abundance of Henslow's sparrows.

We use a data set (`HespData.csv`) consisting of counts of Henslow's sparrows collected along transects at Big Oaks National Wildlife Refuge in southeastern Indiana, USA. The data can be loaded using the code below. Make sure to specify the correct file path within the `setwd(...)` function.

```
setwd("~/")
HespData=read.csv("HespData.csv")
```

The Henslow data set includes a burn unit (`unit`) and transect identifier (`transect`), a unique transect ID (`transectID`), the year data were collected (`year`), the size of the grassland being observed from the transect (`ha`), the number of Henslow's sparrows counted (`counts`), covariates for how many years post-burn the counts occurred (`second.year`...`fourth.year`), and covariates for each transect for the random effect that will be fit (`r.1`...`r.8`). R returns NA for missing values in a .csv file (see below). We want to convert the NAs to 0.

```
head(HespData)
```

```
  unit transect transectID year    ha counts second.year third.year
1    1        a          1 2004 20.68      5          NA         NA
2    1        a          1 2005 20.68     17           1         NA
3    1        a          1 2006 20.68      0          NA          1
4    1        a          1 2007 20.68      2          NA         NA
5    1        b          2 2004  5.04      4          NA         NA
6    1        b          2 2005  5.04     13           1         NA
  fourth.year r.1 r.2 r.3 r.4 r.5 r.6 r.7 r.8
1          NA   1  NA  NA  NA  NA  NA  NA  NA
2          NA   1  NA  NA  NA  NA  NA  NA  NA
3          NA   1  NA  NA  NA  NA  NA  NA  NA
4           1   1  NA  NA  NA  NA  NA  NA  NA
5          NA  NA   1  NA  NA  NA  NA  NA  NA
6          NA  NA   1  NA  NA  NA  NA  NA  NA
```

```
HespData[is.na(HespData)]=0
```

Model Statement

To obtain inference for the unknown average annual abundance of Henslow's sparrows for each burn interval we developed a Bayesian, hierarchical model to fit the data. We used a generalized linear mixed-effects regression model,

$$\begin{aligned} y_{j,t} &\sim \text{Poisson}(A_j \lambda_{j,t}), \\ log(\lambda_{j,t}) &= \boldsymbol{x'}_{j,t}\boldsymbol{\beta} + \eta_j, \\ \boldsymbol{\beta} &\sim \text{Normal}(\boldsymbol{\mu}, \sigma^2 \mathbf{I}), \\ \eta_j &\sim \text{Normal}(0, 1), \end{aligned} \tag{1}$$

where $y_{j,t}$ are the counts of Henslow's sparrows at site $j = 1, ..., 8$ during years $t = 1, ..., T = 4$, A_j is the area of site j, $\lambda_{j,t}$ is the unknown density of Henslow's sparrows at site j in time t and is a function of $\boldsymbol{\theta'} \equiv (\boldsymbol{\beta}, \boldsymbol{\eta})$, and $\boldsymbol{x}_{j,t}$ represents the categorical explanatory variable *summers-post-burn* in site j, year t. The η_j $(j = 1, ..., 8)$ account for differences in densities among sites. We assumed η_j had mean 0 and variance equal to one to reflect the variation in densities among sites. We choose one as the variance because past estimates of densities at Big Oaks were usually between 0 and 2 birds per ha. The mean vector $\boldsymbol{\mu'} = (-5.0, 2.5, 0.2, 0.2)$ for the prior distribution of $\boldsymbol{\beta}$ was obtained using moment matching from density estimates reported in Herkert and Glass (1999). We let $\sigma^2 = 10$ to reflect our uncertainty in $\boldsymbol{\mu}$ because Herkert and Glass (1999) focused on a study site in a different state and during a different time period.

The posterior predictive distributions for the average annual abundance of Henslow's sparrows $[N_{\delta,\boldsymbol{\theta}}|y]$ at the eight sites were derived using the equation:

$$N_{\delta,\boldsymbol{\theta}} = \lim_{\tilde{T}\to\infty} \frac{\sum_{j=1}^{8} \sum_{t=T+1}^{\tilde{T}} a_i \lambda_{j,t}(\delta, \boldsymbol{\theta})}{\tilde{T} - T} \tag{2}$$

Model Fitting

To fit the model described in eq. 1, we wrote a custom MCMC algorithm in R. The algorithm is below.

```
pois.rv.mcmc <- function(y,X,W,betamn,betavar,etamn,etavar,
                         n.mcmc,betatune,etatune){

  #
  #  Perry Williams and Mevin Hooten (20150516)  Last Updated:  20150516
  #  Poisson regression with random effect (GLMM)
  #
  #  EXAMPLE USE:
  #  ------------
  #
  #  hesp.out=pois.rv.mcmc(y,X,W,betamn,betavar,etamn,etavar,
  #       n.mcmc,betatune,etatune)
  #
  #
  #

  ###
  ###  Variables
  ###
```

```
n.burn=round(n.mcmc/10)
X=as.matrix(X)
W=as.matrix(W)
y=as.vector(y)
n=length(y)
p=dim(X)[2]
rho=dim(W)[2]

betasave=matrix(0,p,n.mcmc)
lamsave=matrix(0,n,n.mcmc)
etasave=matrix(0,rho,n.mcmc)

###
###  Starting Values
###

beta=betamn
eta=rep(etamn,rho)
lam=exp(X%*%beta+W%*%eta)
betasave[,1]=beta
lamsave[,1]=lam
etasave[,1]=eta
accept.beta=0
accept.eta=0

###
###  Gibbs Loop
###

for(k in 1:n.mcmc){

  ###
  ### Sample beta
  ###

  betastar=rnorm(p,beta,betatune)
  lamstar=exp(X%*%betastar+W%*%eta)

  mh1=sum(dpois(y,lamstar*A,log=TRUE))+
      sum(dnorm(betastar,betamn,sqrt(betavar),log=TRUE))
  mh2=sum(dpois(y,lam*A,log=TRUE))+
      sum(dnorm(beta,betamn,sqrt(betavar),log=TRUE))

  mhratio=exp(mh1-mh2)
  if(mhratio > runif(1)){
    beta=betastar
    lam=lamstar
    accept.beta=accept.beta+1
  }

  etastar=rnorm(rho,eta,etatune)
  lamstar=exp(X%*%beta +W%*%etastar)

  mh1=sum(dpois(y,lamstar*A,log=TRUE))+
      sum(dnorm(etastar,etamn,sqrt(etavar),log=TRUE))
  mh2=sum(dpois(y,lam*A,log=TRUE))+
      sum(dnorm(eta,etamn,sqrt(etavar),log=TRUE))
```

```
    mhratio=exp(mh1-mh2)
    if(mhratio > runif(1)){
      eta=etastar
      lam=lamstar
      accept.eta=accept.eta+1
    }

    ###
    ### Save Samples
    ###

    betasave[,k]=beta
    lamsave[,k]=lam
    etasave[,k]=eta
  }
  cat("\n");flush.console()

  ###
  ###  Write output
  ###

  list(y=y,X=X,n.mcmc=n.mcmc,betasave=betasave[,n.burn:n.mcmc],
       lamsave=lamsave[,n.burn:n.mcmc],etasave=etasave[,n.burn:n.mcmc],
       accept.beta=accept.beta,accept.eta=accept.eta)
}
```

Data Preparation

The MCMC algorithm (shown above) requires a vector of observations, `y`, the matrix of time-since-burn covariates, `X`, the random effects covariates `W`, hyperparameters for prior distributions, `betamn, betavar, etamn, etavar`, the number of MCMC iterations, `n.mcmc`, and tuning parameters for Metropolis-Hastings updates, `betatune, etatune`. We assign these values in the code, below.

```
y=HespData$counts
n=length(y)
X=cbind(rep(1,n),as.matrix(HespData[,7:9]))
W=as.matrix(HespData[,10:17])
A=HespData$ha
betamn=c(4.95,2.48,0.22,0.18)
betavar=10
etamn=0
etavar=1
n.mcmc=100000
betatune=.1
etatune=.1
```

After assigning the applicable objects, we select a seed, and run the algorithm, keeping track of the time the algorithm takes to run (it might take a few seconds).

```
## Fit the model
set.seed(2016)
sys.time=Sys.time()
hesp.out=pois.rv.mcmc(y,X,W,betamn,betavar,etamn,etavar,n.mcmc,
                      betatune,etatune)
```

```
Sys.time()-sys.time

## Time difference of 4.502345 secs
```

Convergence is shown in Fig. 1, and posterior distributions of $\boldsymbol{\beta}$ are shown in Fig. 2.

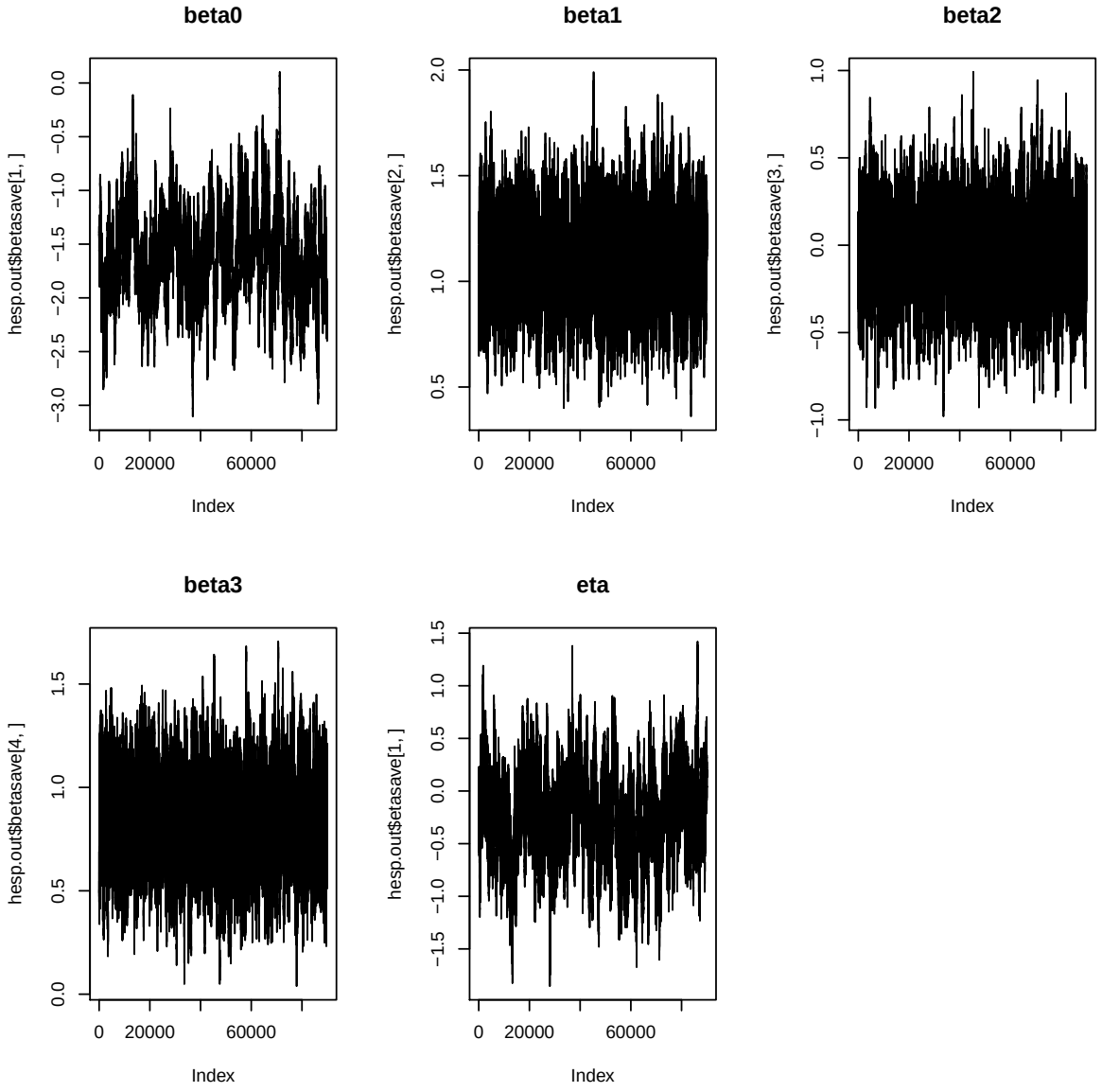

Figure 1: Trace plots to demonstrate convergence

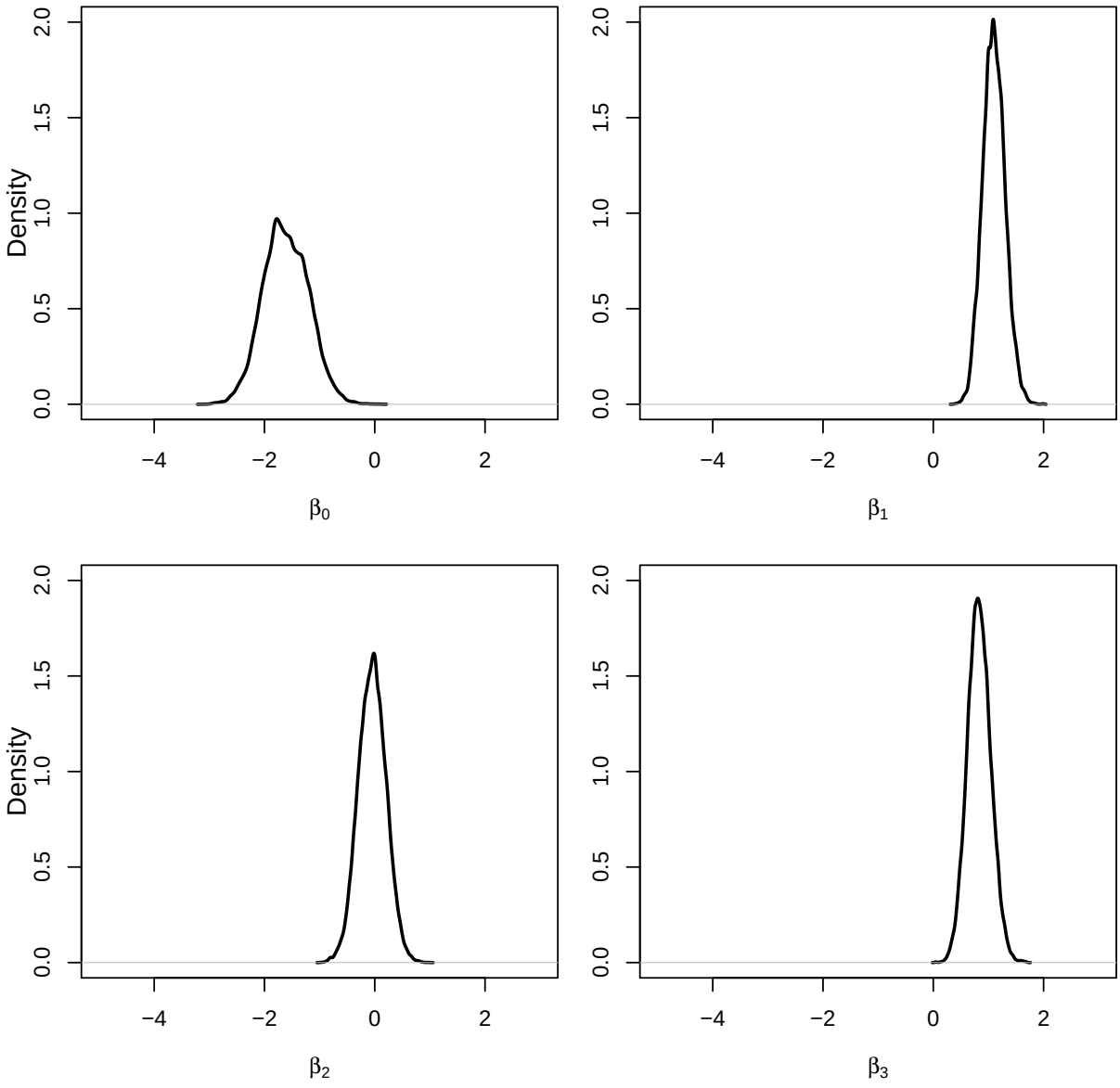

Figure 2: Marginal posterior distributions of $\boldsymbol{\beta}$

From the MCMC iterations, we calculate the derived parameter $N_{\delta,\boldsymbol{\theta}}$ from eq. 2, for each of our possible actions (prescribed fire rotations).

```
###
###  Calculate N(theta,delta)
###

n.iter=length(round(n.mcmc/10):n.mcmc)
T.tilde=1000  # arbitrarily close to \infty
sites=max(HespData$transectID)
a=unique(A)

burn1=burn2=burn3=burn4=matrix(0,nrow=4,ncol=T.tilde)
burn1[1,]=burn2[1,]=burn3[1,]=burn4[1,]=1
burn2[2,]=rep(c(0,1),T.tilde/2)
burn3[2,]=head(rep(c(0,1,0),T.tilde/2),T.tilde)
burn3[3,]=head(rep(c(0,0,1),T.tilde/2),T.tilde)
```

```
burn4[2,]=rep(c(0,1,0,0),T.tilde/4)
burn4[3,]=rep(c(0,0,1,0),T.tilde/4)
burn4[4,]=rep(c(0,0,0,1),T.tilde/4)


derive.ppd=function(n.iter,etasave,T.tilde,sites,X,betasave,a){
  N=matrix(NA,nrow=T.tilde,ncol=n.iter)
  for(i in 1:n.iter){
    eta=matrix(etasave[,i],nrow=T.tilde,ncol=sites,byrow=TRUE)
    xb.tmp=t(X)%*%betasave[,i]
    xb=matrix(rep(xb.tmp,sites),nrow=T.tilde,ncol=sites)
    lambda=exp(xb+eta)
    N[,i]=c(lambda%*%a)
  }
  return(N)
}

N.1.tmp=derive.ppd(n.iter,hesp.out$etasave,
                   T.tilde,sites,burn1,hesp.out$betasave,a)
N.2.tmp=derive.ppd(n.iter,hesp.out$etasave,
                   T.tilde,sites,burn2,hesp.out$betasave,a)
N.3.tmp=derive.ppd(n.iter,hesp.out$etasave,
                   T.tilde,sites,burn3,hesp.out$betasave,a)
N.4.tmp=derive.ppd(n.iter,hesp.out$etasave,
                   T.tilde,sites,burn4,hesp.out$betasave,a)

###
### 20-year cumulative burn history
###

N.1=20*apply(N.1.tmp,2,sum)/T.tilde
N.2=20*apply(N.2.tmp,2,sum)/T.tilde
N.3=20*apply(N.3.tmp,2,sum)/T.tilde
N.4=20*apply(N.4.tmp,2,sum)/T.tilde
```

A plot of the derived posterior distributions for $N_{\delta,\boldsymbol{\theta}}$ are shown in Fig. 3

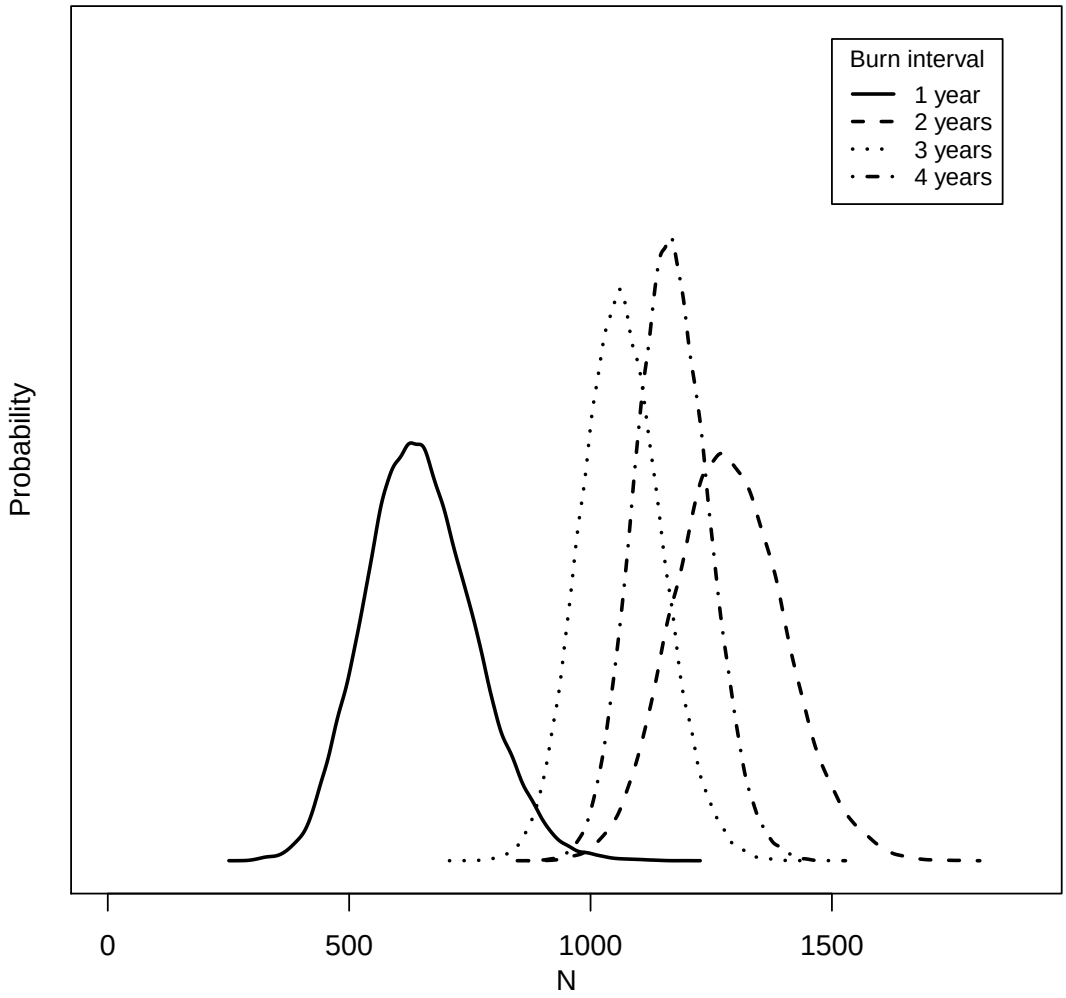

Figure 3: Derived posterior distributions for N(theta,delta) from eq. 2.

Loss Function

To describe loss in terms of the management action, we first developed several axioms the loss function should meet, then developed a quantitative loss function that met all of the axioms. The first axiom was that frequent fire intervals are more costly than infrequent intervals and therefore, all else being equal, frequent fire intervals have higher loss. Second, if cumulative abundance of Henslow's sparrows increases, loss decreases. For our third axiom, we assumed the manager had a dedicated budget for Henslow's sparrow management; if the manager meets the abundance objective (or comes close to meeting the objective), the amount spent is proportionately less important than if the manager was far from meeting the objective. If the manager does not meet the objective, the amount spent is wasted, and has proportionately higher loss than larger abundances. This reflects diminishing marginal returns of saving money as the true cumulative abundance increases. Thus our axiom was: when the cumulative abundance of Henslow's sparrows increases, cost becomes less important. Given these axioms, we developed a simple

quantitative expression for the loss function as

$$L(N,\delta) = \begin{cases} \alpha_0(\delta) + \alpha_1(\delta) N_{\delta,\boldsymbol{\theta}}, & N_{\delta,\boldsymbol{\theta}} < 1835 \\ 0, & N_{\delta,\boldsymbol{\theta}} \geq 1835 \end{cases}$$

(Fig. 4). The loss function is a piecewise function with the first component being a line with negative slope (α_1) and intercept (α_0) that depends on δ and the second component equal to 0 when the abundance is greater than 1835 birds (i.e., the population objective). We chose the intercepts (1, 0.9, 0.8, and 0.7 for 1, 2, 3, and 4 year burn intervals, respectively) so that more frequent burn intervals would have higher loss, and scaled the slope ($\alpha_1(\delta) = \frac{-\alpha_0(\delta)}{1835}$) so the loss would be zero if the average annual population size reached 1835 birds. Thus, cost was incorporated in the differing slopes and intercepts for each action. R code for our loss function is below. The loss functions are shown in Fig. 4

```
###
### Loss function
###

N=0:(92*20)
action=1:4
intercept=c(1,.9,.8,.7)
loss=matrix(NA,nrow=length(intercept),ncol=length(N))
for(i in 1:length(intercept)){
  loss[i,]=-intercept[i]/max(N)*N+intercept[i]
}
```

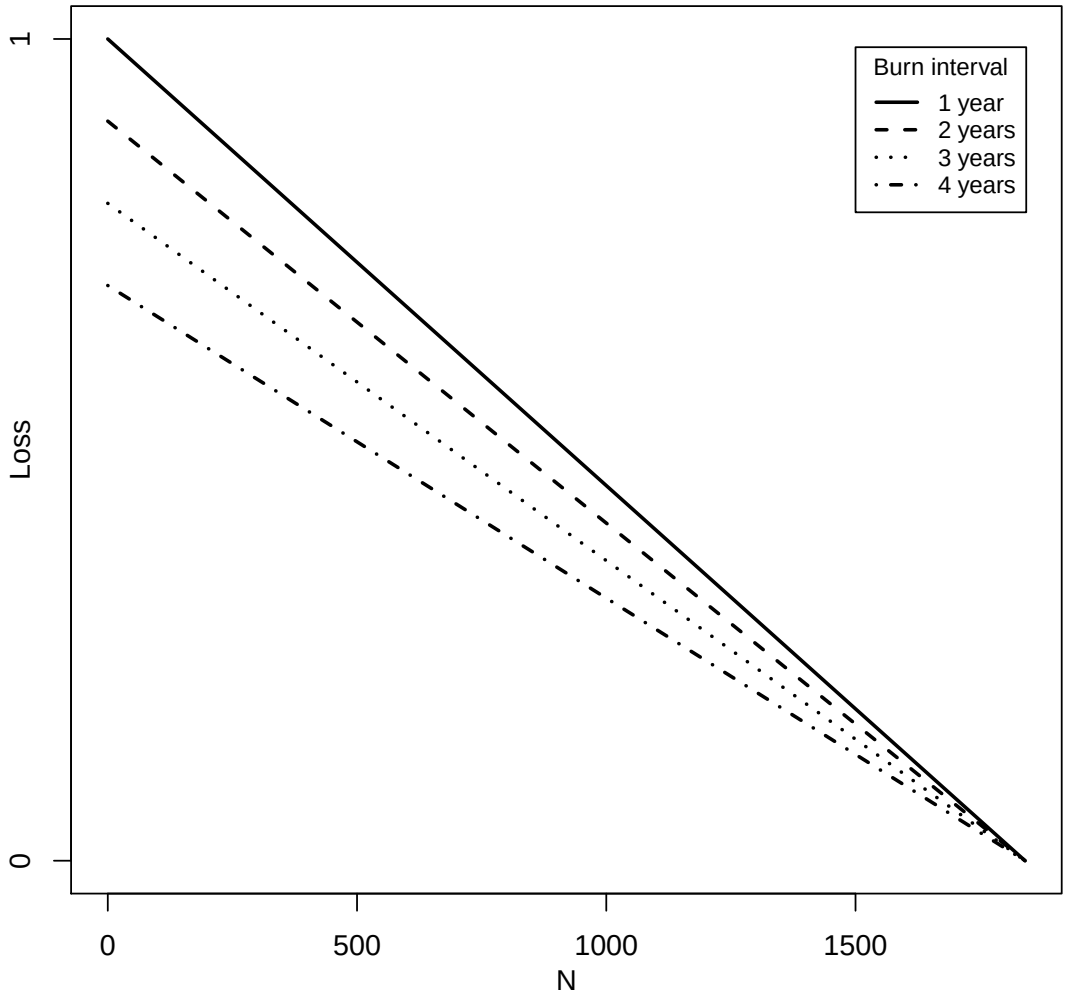

Figure 4: Loss function for Henslow's sparrow management

Bayesian Expected Loss and Optimality

Finally, we identify the optimal prescribed fire rotation for Henslow's sparrow management by calculating the Bayesian expected loss:

$$\rho(a) = E_{\theta|y} L(\theta, a) = \int_{\Theta} L(\theta, a)[\theta|y] d\theta, \tag{3}$$

A plot of the integrand of eq. 3 (i.e., $L(\theta, a)[\theta|y]$) prior to integration is shown in Fig. 5.

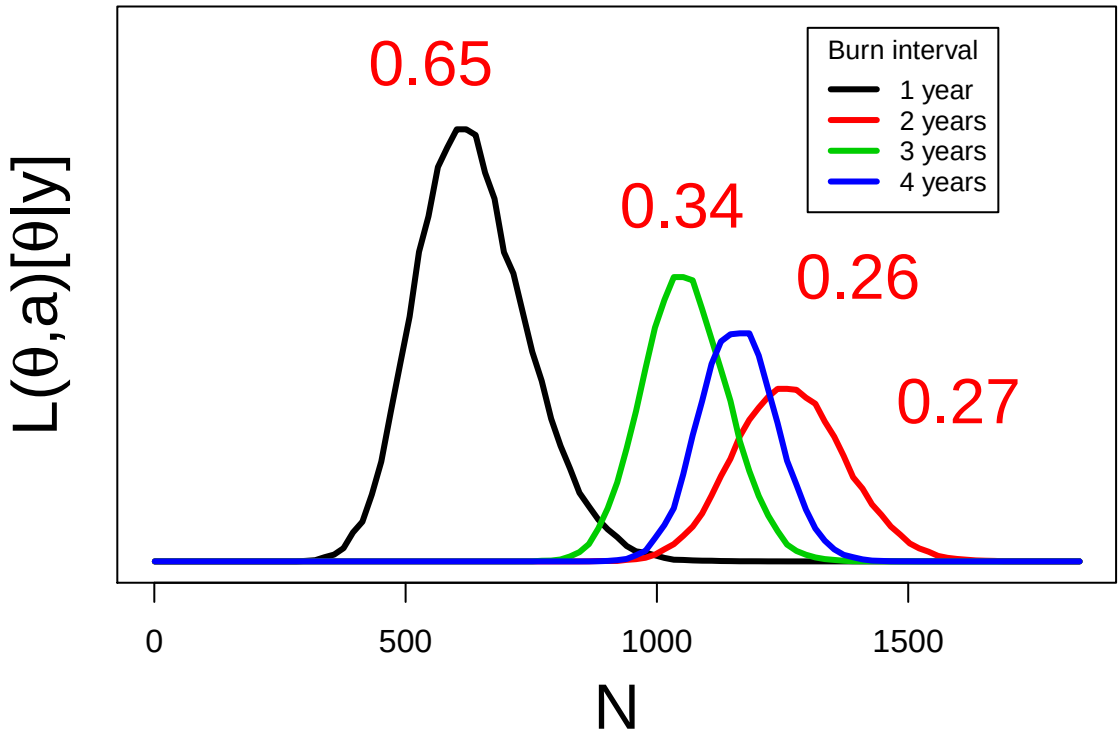

Figure 5: **Lines**: Convolution (i.e., the integrand of eq. 3) of posterior distributions (from Fig. 3) and loss (from Fig. 4) of each potential action. **Numbers**: Bayesian Risk (the result of integration) of each potential action. A four-year burn interval minimizes Bayesian expected loss, and is therefore optimal with respect to the data, model, and loss function.

The R code for calculating Bayesian expected loss, given posterior distributions and a loss function, is straightforward.

```
###
### Calculate Bayesian Expected Loss
### (note: h(dollar-sign)density and loss[i,] have
### to be the same length).
###
N.mat=rbind(N.1,N.2,N.3,N.4)
BayesRisk=numeric(4)
for(i in 1:4){
  h=hist(N.mat[i,],breaks=0:1841,plot=FALSE)
  BayesRisk[i]=sum(h$density*loss[i,])
}
BayesRisk

[1] 0.6478278 0.2702727 0.3372603 0.2557930
```

References

Herkert, J.R., and W.D. Glass. 1999. Henslow's sparrow response to prescribed fire in an Illinois prairie remnant. Studies in Avian Biology 19:160:164.

Williams, P.J., and M.B. Hooten. 2016. Combining statistical inference and decisions in ecology. Ecological Applications. DOI: 10.1890/15-1593.1.