Chapter 1

Introduction

Chapter 2

GLMS and WinBUGS

Chapter 3

Estimating the Size of a Closed Population

In this chapter we will consider ordinary capture-recapture (CR) models for estimating population size in closed populations. We will see that such models are closely related to binomial (or logistic) regression type models. In fact, when N is known, they are precisely such models. We consider some important extensions of ordinary closed population models that accommodate various types of "individual effects" — either in the form of explicit covariates (sex, age, body mass) or unstructured "heterogeneity" in the form of an individual random effect. In general, these models are variations of generalized linear or generalized linear mixed models (GLMMs). Because of the paramount importance of this concept, we focus mainly on fairly simple models in which the observations are individual encounter frequencies, y_i = the number of encounters of individual i out of K replicate samples of the population which, for the models we consider here, is the outcome of a binomial random variable. Along the way, we consider the spatial context of capture-recapture data and models and demonstrate that density cannot be formally estimated when spatial information is ignored. We also review some of the informal methods of estimating density using CR methods, and consider some of their limitations. We will be exposed to our first primitive spatial capture-recapture models which arise as relatively minor variations of so-called "individual covariate models" (of the Huggins (1989) and Alho (1990) variety). In a sense, the point of this chapter is to establish that linkage in a direct and concise manner beginning with the basic "Model M0" and extensions of that model to include individual heterogeneity and also individual covariates. A special type of individual covariate models is distance sampling, which could be thought of as the most primitive spatial capture-recapture model. In later chapters we further develop and extend ideas introduced in this chapter.

We emphasize Bayesian analysis of capture-recapture models and we accomplish this using a method related to classical "data augmentation" from the statistics literature Tanner and Wong (e.g., 1987)). This is a general concept in statistics but, in the context of capture-recapture models where N is unknown, it has a consistent implementation across classes of capture-recapture models and one that is really convenient from the standpoint of doing MCMC (Royle et al., 2007). We use data augmentation throughout this book and thus emphasize its conceptual and technical origins and demonstrate applications to closed population models. We refer the reader to Kery and Schaub (2011, ch. 6) for an accessible and complimentary development of ordinary closed population models.

3.1 The Simplest Closed Population Model: Model M0

We suppose that there exists a population of N individuals which we subject to repeated sampling, say over K nights, where individuals are captured, marked, and subsequently recaptured. We suppose that individual encounter histories are obtained, and these are of the form of a sequence of 0's and 1's indicating capture ($y = 1$) or not ($y = 0$) during any sampling occasion ("sample"). As an example, suppose $K = 5$ sampling occasions, then an individual captured during sample 2 and 3 but not otherwise would have an encounter history of the form $\mathbf{y} = (0, 1, 1, 0, 0)$. Thus, the observation $\mathbf{y}_i$ for each individual (i) is a vector having elements denoted by y_{ik} for $k = 1, 2, .., K$. Usually this is organized as a row of a matrix with elements y_{ik}, see Table 3.1. Except where noted explicitly, we suppose that observations are independent within individuals and among individuals. Formally, this allows us to say that y_{ik} are Bernoulli random variables and we may write $y_{ik} \sim \text{Bern}(p)$. Consequently, for this very simple model in which p is in fact constant, then we can declare that the individual encounter frequencies (total captures), $y_i = \sum_k y_{ik}$, have a binomial distribution based on a sample of size K. That is

$$y_i = \sum_k y_{ik} \sim \text{Bin}(p, K)$$

for every individual in the population. This is a remarkably simple model that forms the cornerstone of almost all of classical capture-recapture models, including most spatial capture-recapture models discussed throughout this book. Evidently, the basic capture-recapture model structure is precisely a simplistic version of a logistic-regression model with only an intercept term ($\text{logit}(p) = \text{constant}$). To say that all capture-recapture models are just logistic regressions is only slightly inaccurate. In fact, we are proceeding here "conditional on N", i.e., as if we knew N. In practice we don't, of course, and that is kind of the point of capture-recapture models as estimating N is the central objective. But, by proceeding conditional on N, we can specify a simple model and then deal with the fact that N is unknown using standard methods that you are already familiar with (i.e., GLMs - see chapter 2).

Table 3.1: a capture-recapture data set with $n = 6$ observed individuals and $K = 5$ samples.

	Sample occasion					
indiv i	1	2	3	4	5	y_i
1	1	0	0	1	0	2
2	0	1	0	0	1	2
3	1	0	0	1	0	2
4	1	0	1	0	1	3
5	0	1	0	0	0	1
$n = 6$	1	0	0	0	0	1

Assuming individuals of the population are observed independently, the joint probability distribution of the observations is the product of N binomials

$$\begin{aligned}\Pr(y_1, \ldots, y_N | p) &= \prod_{i=1}^{N} \text{Bin}(y_i | K, p) \\ &= \prod_{k=0}^{K} \pi(k)^{n_k}\end{aligned}$$

where $\pi(k) = \text{Bin}(k|K, p)$ and where $n_k = \sum_{i=1}^{N} I(y_i = k)$ denotes the number of individuals captured k times in K surveys. We emphasize that this is conditional on N, in which case we get to observe the $y = 0$ observations and the resulting data are just *iid* binomial counts. Because this is a binomial regression model of the variety described in chapter 2, fitting this model using a BUGS engine poses no difficulty.

The essential problem in capture-recapture, however, is that N is not known because the number of uncaptured/missing individuals (i.e., those in the zero cell that occur with probability $\pi(0)$) is unknown. Consequently, the observed capture frequencies n_k are no longer independent. Instead, their joint distribution is multinomial (e.g., see Illian (2008, p. xyz)):

$$n_1, n_2, \ldots, n_K \sim \text{Multin}(N, \pi(1), \pi(2), \ldots, \pi(K)) \tag{3.1}$$

Note that in our notation the number of uncaptured/missing individuals is denoted by $n_0 = N - n$, where $n = \sum_{k=1}^{K} n_k$ denotes the total number of distinct individuals seen in the K samples.

To fit the model in which N is *unknown*, we can regard N as a parameter and maximize the multinomial likelihood directly. While direct likelihood analysis of the multinomial model is straightforward, that does not prove to be too useful in practice because we seldom are concerned with models for the aggregated encounter history frequencies. In many instances, including for spatial capture-recapture (SCR) models, we require a formulation of the model that can accommodate individual level covariates which we address subsequently in this chapter.

3.1.1 The Spatial Context of Capture-Recapture

A common assumption made is that of population "closure" which is really just a colloquial way of saying (in part) the Bernoulli assumptions stated explicitly above. In the biological context, closure means, strictly, no additions or subtractions from the population during study. This is manifest by the statement that the encounters are independent and identically distributed (iid) Bernoulli trials. In practice, closure is usually interpreted by the manner in which potential violations of that assumption arise. In particular, two important elements of the closure assumption are "demographic" and "geographic" closure. If an individual dies then subsequent values of y_{ik} are clearly no longer Bernoulli trials with the same parameter p. If there is no mortality or recruitment in the population, then we say that demographic closure is satisfied. Similarly, animals may emigrate or immigrate. If they do not, then geographic closure is satisfied. Sometimes a distinction is made between temporary and permanent emigration or immigration. That is a relevant distinction in spatial capture-recapture models, because SCR models explicitly accommodate "temporary emigration" of a certain type, due to individuals moving about their home range. The demographic closure assumption can also be relaxed using SCR models, but we will save that discussion for chapter 4.

3.1.2 Conditional likelihood

We saw that a basic closed population model is a simple logistic regression model if N is known and, when N is unknown, the model is multinomial with index or sample size parameter N. This multinomial model, being conditional on N, is sometimes referred to as the "joint likelihood" the "full likelihood" or the "unconditional likelihood" (or model in place of likelihood). This formulation differs from the so-called "conditional likelihood" approach in which the likelihood of the observed encounter histories is devised conditional on the event that an individual is captured at least once. To construct this likelihood, we have to recognize that individuals appear or not in the sample based on the value of the random variable y_i, that is, we capture them if and only if $y_i > 0$. The observation model is therefore based on $\Pr(y|y > 0)$. For the simple case of Model M0, the resulting conditional distribution is a "zero truncated" binomial distribution which accounts for the fact that we cannot observe the value $y = 0$ in the data set (see Royle and Dorazio, 2008, section XYZ). Both the conditional or unconditional models are legitimate modes of analysis in all capture-recapture types of studies, and they provide equally valid descriptions of the data and for many practical purposes provide equivalent inferences, at least in large sample sizes (Sanathanan, 1972).

In this book we emphasize Bayesian analysis of capture-recapture models using data augmentation (discussed subsequently), which produces yet a third distinct formulation of capture recapture-models based on the zero-*inflated* binomial distribution that we describe in the next section. Thus, there are 3 distinct formulations of the model – or models of analysis – for analyzing all

Mode of analysis	parameters in model	statistical model
Joint likelihood	p, N	multinomial with index N
Conditional likelihood	p	zero-truncated binomial
Data augmentation	p, ψ	zero-inflated binomial

Table 3.2: Modes of analysis of capture-recapture models.

capture-recapture models based on the (1) binomial model for the joint or unconditional specification; (2) zero-truncated binomial that arises "conditional on n"; and (3) the zero-inflated binomial that arises under data augmentation. Each formulation has a distinct complement of model parameters (shown in Table 3.2 for Model M0).

3.2 Data Augmentation

We consider a method of analyzing closed population models using data augmentation (DA) which is useful for Bayesian analysis and, in particular, analysis of models using the various BUGS engines and other software. Data augmentation is a general statistical concept that is widely used in statistics in many different settings. The classical reference is Tanner and Wong (1987) but see also Liu and Wu (1999). Data augmentation can be adapted to provide a very generic framework for Bayesian analysis of capture-recapture models with unknown N. This idea was introduced for closed populations by Royle et al. (2007), and has subsequently been applied to a number of different contexts including individual covariate models (Royle, 2009), open population models (Royle and Dorazio, 2008, 2010; Gardner et al., 2010), spatial capture-recapture models (Royle and Young, 2008; Royle, 2010; Gardner, 2009), and many others.

Conceptually, data augmentation takes the data you wish you had - that is, the data set with N rows - the known-N data set - and embeds that data set into a larger data set having $M > N$ rows. [1] It is always possible, in practice, to choose M pretty easily for a given problem and context. Then, under data augmentation, analysis is focused on the "augmented data set." That is, we analyze the bigger data set - the one having M rows - with an appropriate model that accounts for the augmentation. Inference is focused directly on estimating the proportion $\psi = E[N]/M$, instead of directly on N, where ψ is the "data augmentation parameter."

3.2.1 DA links occupancy models and closed population models

We provide a heuristic description of data augmentation based on the close correspondence between so-called "occupancy" models and closed population

[1] RC: Might be just me, but I find that formulation a little confusing... I think it's the 'data you wish you had because that's effectively data you don't have. I think it might be easier to grasp if this were explained with the data you do have - based on n.

models following Royle and Dorazio (2008, sec. xyz).

In occupancy models (MacKenzie et al., 2002; Tyre et al., 2003) the sampling situation is that M sites, or patches, are sampled multiple times to assess whether a species occurs at each site. This yields encounter data such as that illustrated in the left panel of Table 3.3. The important problem is that a species may occur at a site, but go undetected, yielding the "all-zero" encounter histories which are observed. However, some of the all-zeros may well correspond to sites where the species in fact *does not* occur. Thus, while the zeros are observed, there are too many of them and, in a sense, the inference problem is to allocate the zeros into "structural" (fixed) and "sampling" (or stochastic) zeros. More formally, inference is focused on the parameter ψ, the probability that a site is occupied. In contrast, in classical closed population studies, we observe a data set as in the middle panel of Table 3.3 where *no* zeros are observed. The inference problem is, essentially, to estimate how many sampling zeros there are - or should be - in a "complete" data set. The inference objective (how many sampling zeros?) is precisely the same for both types of problems if an upper limit M is specified for the closed population model. The only distinction being that, in occupancy models, M is set by design (i.e., the number of sites to visit) whereas a natural choice of M for capture-recapture models may not be obvious. However, by assuming a uniform prior for N on the integers $[0, M]$, this upper bound is induced (Royle et al., 2007). Then, one can analyze capture-recapture models by adding $M - n$ all-zero encounter histories to the data set and regarding the augmented data set, essentially, as a site-occupancy data set.

Thus, the heuristic motivation of data augmentation is to fix the size of the data set by adding *too many* all-zero encounter histories to create the data set shown in the right panel of Table 3.3 - and then analyze the augmented data set using an occupancy type model which includes both "unoccupied sites" as well as "occupied sites" at which detections did not occur. We call these $M - n$ all-zero histories "potential individuals" because they exist to be recruited (in a non-biological sense) into the population, for example during an analysis by MCMC.

To analyze the augmented data set, we recognize that it is a zero-inflated version of the known-N data set. That is, some of the augmented all-zeros are sampling zeros (corresponding to actual individuals that were missed) and some are "structural" zeros, which do not correspond to individuals in the population. For a basic closed-population model, the resulting likelihood under data augmentation - that is, for the data set of size M – is a simple zero-inflated binomial likelihood. The zero-inflated binomial model can be described "hierarchically", by introducing a set of binary latent variables, $z_1, z_2, \ldots, z_M$, to indicate whether each individual i is $(z_i = 1)$ or is not $(z_i = 0)$ a member of the population of N individuals exposed to sampling. We assume that $z_i \sim \text{Bern}(\psi)$ where ψ is the probability that an individual in the data set of size M is a member of the sampled population - in the sense that $1 - \psi$ is the probability of realizing a "structural zero" in the augmented data set. The zero-inflated binomial model which arises under data augmentation can be formally expressed bythe following set of assumptions:

$$\begin{aligned} y_i|z_i = 1 &\sim \text{Bin}(K, p) \\ y_i|z_i = 0 &\sim \delta(0) \\ z_i &\overset{iid}{\sim} \text{Bern}(\psi) \\ \psi &\sim \text{Unif}(0, 1) \\ p &\sim \text{Unif}(0, 1) \end{aligned}$$

for $i = 1, \ldots, M$, where $\delta(0)$ is a point mass at $y = 0$.

We note that N is no longer an explicit parameter of this model. Instead, we estimate ψ and functions of the latent variables. In particular, under the assumptions of the zero-inflated model, $z_i \overset{iid}{\sim} \text{Bern}(\psi)$; therefore, N is a function of these latent variables:

$$N = \sum_{i=1}^{M} z_i.$$

Further, we note that the latent z_i parameters can be removed from the model by integration, in which case the joint probability of the data is

$$\Pr(y_1, \ldots, y_M | p, \psi) = \prod_{i=1}^{M} \psi \text{Bin}(y_i | K, p) + I(y_i = 0)(1 - \psi) \tag{3.2}$$

Which can be maximized directly to obtain the MLEs of the structural parameters ψ and p or those of other more complex models (e.g., see Royle, 2006). We could estimate these parameters and then use them to obtain an estimator of N using the so-called "Best unbiased predictor" (see Royle and Dorazio, 2011).

3.2.2 Model M_0 in BUGS

For model M_0 in which we can aggregate the encounter data to individual-specific encounter frequencies, the augmented data are given by the vector of frequencies $(y_1, \ldots, y_n, 0, 0, \ldots, 0)$. The zero-inflated model of the augmented data combines the model of the latent variables, $z_i \sim \text{Bern}(\psi)$ with the conditional-on-z binomial model:

$$\begin{aligned} y_i|z_i = 0 &\sim \delta(0) \\ y_i|z_i = 1 &\sim \text{Bin}(K, p) \end{aligned}$$

It is convenient to express the conditional-on-z observation model concisely as:

$$y_i|z_i \sim \text{Bin}(K, p z_i)$$

Thus, if $z_i = 0$ then the success probability of the binomial distribution is identically 0 whereas, if $z_i = 1$, then the success probability is p. This is useful in describing the model in the **BUGS** language, as shown below. Note the last line of the model specification here provides the expression for computing N from the data augmentation variables z_i.

Table 3.3: Hypothetical occupancy data set (left), capture-recapture data in standard form (center), and capture-recapture data augmented with all-zero capture histories (right).

Occupancy data				Capture-recapture				Augmented C-R			
site	k=1	k=2	k=3	ind	k=1	k=2	k=3	ind	k=1	k=2	k=3
1	0	1	0	1	0	1	0	1	0	1	0
2	1	0	1	2	1	0	1	2	1	0	1
3	0	1	0	.	0	1	0	3	1	0	1
4	1	0	1	.	1	0	1	4	1	0	1
5	0	1	1	.	0	1	1	5	1	0	1
.	0	1	1	.	0	1	1	.	0	1	1
.	1	1	1	.	1	1	1	.	0	1	1
.	1	1	1	.	1	1	1	.	1	1	1
	1	1	1	.	1	1	1	.	1	1	1
n	1	1	1	n	1	1	1	n	1	1	1
.	0	0	0					.	0	0	0
.	0	0	0					.	0	0	0
	0	0	0						0	0	0
	0	0	0						0	0	0
	0	0	0						0	0	0
	0	0	0					N	0	0	0
.	0	0	0					.	0	0	0
.	0	0	0						0	0	0
M	0	0	0					.	0	0	0
								.	.	.	.
								.	.	.	.
								.	.	.	.
								M	0	0	0

```
p  ~ dunif(0,1)
psi~dunif(0,1)

# nind = number of individuals captured at least once
#   nz = number of uncaptured individuals added for PX-DA
for(i in 1:(nind+nz)) {
    z[i]~dbern(psi)
   mu[i]<-z[i]*p
    y[i]~dbin(mu[i],K)
 }

N<-sum(z[1:(nind+nz)])
```

Specification of a more general model in terms of the individual encounter observations y_{ik} is not much more difficult than for the individual encounter frequencies. We define the observation model by a double loop and change the indexing of things accordingly, i.e.,

```
for(i in 1:(nind+nz)) {
    z[i]~dbern(psi)
  for(k in 1:K){
      mu[i,k]<-z[i]*p
      y[i,k]~dbin(mu[i,k],1)
  }
}
```

In this manner, it is straightforward to incorporate covariates on p (see discussion of this below and also chapt. 8 (REF XYZ) and consider other extensions.

3.2.3 Formal development of data augmentation

Use of DA for solving inference problems with unknown N can be justified as originating from the choice of uniform prior on N. The $\mathrm{Unif}(0, M)$ prior for N is innocuous in the sense that the posterior associated with this prior is equal to the likelihood for sufficiently large M. One way of inducing the $\mathrm{Unif}(0, M)$ prior on N is by assuming the following hierarchical prior:

$$\begin{aligned} N &\sim \mathrm{Bin}(M, \psi) \\ \psi &\sim \mathrm{Unif}(0, 1) \end{aligned} \tag{3.3}$$

which includes a new model parameter ψ. This parameter denotes the probability that an individual in the super-population of size M is a member of the population of N individuals exposed to sampling. The model assumptions, specifically the multinomial model (eq. XYZ) and eq. 3.3, may be combined to yield a reparameterization of the conventional model that is appropriate for the augmented data set of known size M:

$$(n_1, n_2, \ldots, n_K) \sim \mathrm{Multin}(M, \psi\pi(1), \psi\pi(2), \ldots, \psi\pi(K)) \tag{3.4}$$

This arises by removing N from Eq. multinomial XYZ by integrating over the binomial prior distribution for N. Thus, the models we analyze under data augmentation arise formally by removing the parameter N from the ordinary model - the model conditional on N - by integrating over a binomial prior distribution for N.

Note that the $M - n$ unobserved individuals in the augmented data set have probability $\psi\pi(0) + (1 - \psi)$, indicating that these unobserved individuals are a mixture of individuals that are sampling zeros ($\psi\pi_0$, and belong to the population of size N) and others that are "structural zeros" (occurring in the augmented data set with probability $1-\psi$). In Eq. 3.4 N has been eliminated as a formal parameter of the model by marginalization (integration) and replaced with the new parameter ψ, which we will call the "data augmentation parameter." However, the full likelihood containing both N and ψ can be analyzed (see Royle et al., 2007).

3.2.4 Remarks on Data Augmentation

Data augmentation may seem like a strange and mysterious black-box, and likely it is unfamiliar to most people even those with extensive experience with capture-recapture models. However, it really is a formal reparameterization of capture-recapture models in which N is removed from the ordinary (conditional-on-N) model by integration. In the case of Model M0, data augmentation produces the zero-inflated binomial which is distinct from the original observation model, but only in the sense that it embodies, explicitly, the Unif$(0, M)$ prior for N. Choice of M might be cause for some concern related to potential sensitivity to choice of M. The guiding principle is that it should be chosen large enough so that the posterior for N is not truncated, but no larger because large values entail more computational burden. It seems likely that the properties of the Markov chains should be affected by M and so some optimality might exist (Gopalaswamy, 2012), as in occupancy models (Mackenzie and Royle, 2005). Formal analysis of this is required.

We emphasize the motivation for data augmentation being that it produces a data set of fixed size, so that the parameter dimension in any capture-recapture model is also fixed. As a result, MCMC is a relatively simple proposition using standard Gibbs Sampling. Consider the simplest context - analyzing Model M0 using the occupancy model. In this case, DA converts Model M0 to a basic occupancy model and the parameters p and ψ have known full-conditional distributions (in fact, beta distributions) that can be sampled from directly. Furthermore, the data augmentation variables - the latent data augmentation variables z, can be sampled from Bernoulli full conditionals. MCMC is not too much more difficult for complicated models - sometimes the hyperparameters need to be sampled using a Metropolis-Hastings step, but nothing more sophisticated than that is required.

There are other approaches to analyzing models with unknown N, using reversible jump MCMC (RJMCMC) or other so-called "trans-dimensional" (TD)

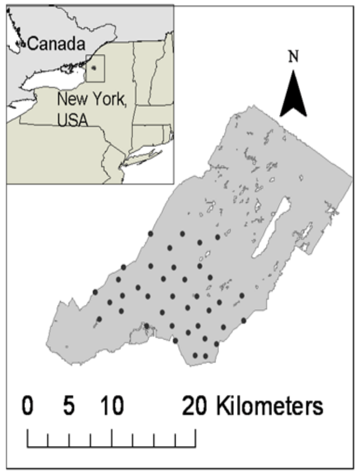

Figure 3.1: Fort Drum study area and hair snare locations.

algorithms[2] (Durbin and Elston, 2012; King, missing; Schofield and Barker, missing). What distinguishes DA from RJMCMC and related TD methods is that DA is used to create a distinctly new model that is unconditional on N and we (usually) analyze the unconditional model. The various TD/RJMCMC approaches seek to analyze the conditional-on-N model in which the dimensional of the parameter space is a variable function of N. TD/RJMCMC approaches might appear to have the advantage that one can model N explicitly or consider alternative priors for N. However, despite that N is removed as an explicit parameter in DA, it is possible to develop hierarchical models that involve structure on N (Converse and Royle, 2010; **?**) which we consider in chapt. XYZ.

3.2.5 Example: Black Bear Study on Fort Drum

To illustrate the analysis of Model M0 using data augmentation, we use a data set collected at Fort Drum Military Installation in upstate New York by the Department of Defense, Cornell University and colleagues. These data have been analyzed in various forms by Gardner (2009); Gardner et al. (2010), and Wegan (missing). The specific data used here are encounter histories on 47 individuals obtained from an array of 38 baited "hair snares" (Fig. 3.1) during June and July 2006. Barbed wire traps were baited and checked for hair samples each week for eight weeks, thus we have $K = 8$ sample intervals. The data are provided on the Web Supplement and the analysis can be set up and run as follows. Here, the data were augmented with $M - n = 128$ ($M = 175$) all-zero encounter histories.

```
# Consider adding comments to your code.
```

[2]Look these citeations up in Royle-Dorazio EURING paper

```
## Good idea. This will be done in final draft
trapmat<-read.csv("FDtrapmat.csv")
bearArray<-source("FDbeararray.R")$value
nind<-dim(bearArray)[1]
K<-dim(bearArray)[3]
ntraps<-dim(bearArray)[2]

M=175
nz<-M-nind

Xaug <- array(0, dim=c(M,ntraps,K))
Xaug[1:nind,,]<-bearArray
y<- apply(Xaug,c(1,3),sum)
y[y>1]<-1
ytot<-apply(y,1,sum)   # total encounters out of K
```

Note that the raw data, $\mathbf{y}$, is an $M \times K$ array of individual encounter events (i.e., $y_{ik} = 1$ if individual i was encountered in any trap and 0 otherwise). For $i = 48, ..., 175$, y_{ik}=0 as these are augmented observations. For Model M0 it is sufficient to reduce the data to individual encounter frequencies which we have labeled `ytot` above. The BUGS model file along with commands to fit the model are as follows:

```
set.seed(2013)                # to obtain the same results each time
data0<-list(y=y,M=M,K=K)
params0<-list('psi','p','N')
zst=c(rep(1,nind),rbinom(M-nind, 1, .5))
inits =  function() {list(z=zst, psi=runif(1), p=runif(1)) }

cat("
model {

psi~dunif(0, 1)
p~dunif(0,1)

for (i in 1:M){
   z[i]~dbern(psi)
   for(k in 1:K){
      tmp[i,k]<-p*z[i]
      y[i,k]~dbin(tmp[i,k],1)
       }
       }
N<-sum(z[1:M])
}
",file="modelM0.txt")

fit0 = bugs(data0, inits, params0, model.file="modelM0.txt",
       n.chains=3, n.iter=2000, n.burnin=1000, n.thin=1,
       debug=TRUE,working.directory=getwd())
```

The posterior summary statistics from this analysis are as follows:

```
> print(fit0,digits=2)
Inference for Bugs model at "modelM0.txt", fit using WinBUGS,
 3 chains, each with 2000 iterations (first 1000 discarded)
 n.sims = 3000 iterations saved
           mean    sd   2.5%    25%    50%    75%  97.5% Rhat n.eff
psi        0.29  0.04   0.22   0.26   0.29   0.31   0.36    1  3000
p          0.30  0.03   0.25   0.28   0.30   0.32   0.35    1  3000
N         49.94  1.99  47.00  48.00  50.00  51.00  54.00    1  3000
deviance 489.05 11.28 471.00 480.45 488.80 495.40 513.70    1  3000

[.. some output deleted ...]
```

WinBUGS did well in choosing an MCMC algorithm for this model – we have $\hat{R} = 1$ for each parameter, and an effective sample size of 3000, equal to the total number of posterior samples. We see that the posterior mean of N under this model is 49.94 and a 95% posterior interval is $(48, 54)$. We revisit these data later in the context of more complex models.

In order to obtain an estimate of density, D, we need an area to associate with the estimate of N, and commonly used procedures to conjure up such an area include buffering the trap array by the home range radius, often estimated by the mean maximum distance moved (MMDM)[3], 1/2 MMDM (Dice, 1938) or directly from telemetry data (REF XXX NEED REF HERE XXXXX). Typically, the trap array is defined by the convex hull around the trap locations, and this is what we applied a buffer to. We computed the buffer by using an estimate of the mean female home range radius (2.19 km) estimated from telemetry studies (Bales et al., 2005) instead of using an estimate based on our relatively more sparse recapture data[4]. For the Fort Drum study, the convex hull has area 157.135 km^2, and the buffered convex hull has area 277.011 km^2. To create this we used functions contained in the **R** package `rgeos` and created a utility function `bcharea` which is in our **R** package `scrbook`. The commands are as follows:

```
library("rgeos")

bcharea<-function(buff,traplocs){
p1<-Polygon(rbind(traplocs,traplocs[1,]))
p2<-Polygons(list(p1=p1),ID=1)
p3<-SpatialPolygons(list(p2=p2))
p1ch<-gConvexHull(p3)
 bp1<-(gBuffer(p1ch, width=buff))
 plot(bp1, col='gray')
 plot(p1ch, border='black', lwd=2, add=TRUE)
 gArea(bp1)
}
```

[3] really MMDM? How can this be an estimate of the home range radius? Reference for this?

[4] BETH: Why?

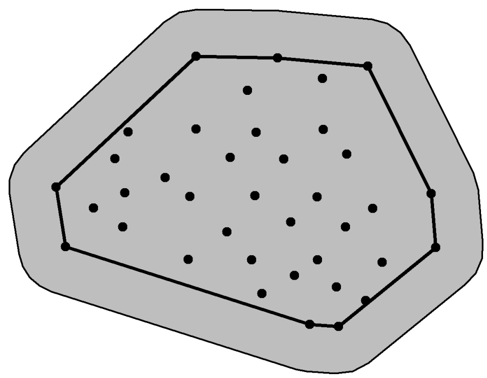

Figure 3.2: buffered convex hull of the bear hair snare array

```
bcharea(2.19,traplocs=trapmat)
```

The resulting buffered convex hull is shown in Fig. 3.2.

To conjure up a density estimate under model M_0, we compute the appropriate posterior summary of N and the prescribed area (277.011 km^2):

```
> summary(fit0$sims.list$N/277.011)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.1697  0.1733  0.1805  0.1803  0.1841  0.2130

> quantile(fit0$sims.list$N/277.011,c(0.025,0.975))
     2.5%     97.5%
0.1696684 0.1949381
```

which yields a density estimate of about 0.18 ind/km^2, and a 95% Bayesian confidence interval of $(0.170, 0.195)$.

The obvious limitation of this estimate and, indeed, of the whole process, is that our choice of "area" is completely subjective - which area should we use? MMDM? One-half MMDM? Estimated from telemetry data? And, furthermore, how certain are we of this area? Can we quantify our uncertainty about this quantity? More important, what exactly is the meaning of this area and, in this context, how do we gauge bias and/or variance of "estimators" of it? (i.e., what is it estimating?).

3.3 Temporally varying and behavioral effects

The purpose of this chapter is mainly to emphasize the central importance of the binomial model in capture-recapture and so we have considered models for individual encounter frequencies - the number of times individuals are captured out of K samples. Sometimes it is not acceptable to aggregate the encounter data for each individual – such as when encounter probability varies over time among samples. A type of time-varying response that seems relevant in most capture-recapture studies is "effort" such as amount of search time, number of observers, or trap effort or when p depends on date (Kéry et al., 2010; Gardner et al., 2010). A common situation is that in which there exists a "behavioral response" to trapping (even if the animal is not physically trapped).

Behavioral response is an important concept in carnivore studies because individuals might learn to come to baited traps or avoid traps due to trauma related to being encountered. There are a number of ways to parameterize a behavioral response to encounter. The distinction between persistent and ephemeral was made by Yang and Chao (2005) who considered a general behavioral response model of the form:

$$\text{logit}(p_{ik}) = \alpha_0 + \alpha_1 * y_{i,k-1} + \alpha_2 x_{ik}$$

where x_{ik} is a covariate indicator variable of previous capture (i.e., $x_{ik} = 1$ if captured in any previous period). Therefore, encounter probability changes depending on whether an individual was captured in the immediate previous period (ephemeral behavioral response) or in any previous period (persistent behavioral response). The former probably models a behavioral response due to individuals moving around their territory relatively slowly over time and the latter probably accommodates trap happiness due to baiting or shyness due to trauma. In spatial capture-recapture models it makes sense to consider a local behavioral response that is trap-specific (**?**) - that is, the encounter probability is modified for individual traps depending on previous capture in specific traps.

Models with temporal effects are easy to describe in the **BUGS** language and analyze and we provide a number of examples in chapt. 8. XXXXX ?? XXXXX

3.4 Models with individual heterogeneity

Here we consider models with individual-specific encounter probability parameters, say p_i, which we model according to some probability distribution, $g(\theta)$. We denote this basic model assumption as $p_i \sim g(\theta)$. This type of model is similar in concept to extending a GLM to a GLMM but in the capture-recapture context N is unknown. The basic class of models is often referred to as "Model M_h" but really this is a broad class of models, each being distinguished by the specific distribution assumed for p_i. There are many different varieties of Model M_h including parametric and various putatively non-parametric approaches (Burnham and Overton, 1978; Norris III and Pollock, 1996; Pledger, 2000). One important practical matter is that estimates of N can be extremely sensitive to the choice of heterogeneity model (Fienberg et al., 1999; Dorazio and Royle, 2003; Link, 2003). Indeed, Link (2003) showed that in some cases it's possible to find models that yield precisely the same expected data, yet produce wildly different estimates of N. In that sense, N for most practical purposes is not identifiable across classes of mixture models, and this should be understood before fitting any such model. One solution to this problem is to seek to model explicit factors that contribute to heterogeneity, e.g., using individual covariate models (See 3.5 below). Indeed, spatial capture-recapture models seek to do just that, by modeling heterogeneity due to the spatial organization of individuals in relation to traps or other encounter mechanism. For additional background and applications of Model M_h see Royle and Dorazio (2008, chapt. 6) and Kery and Schaub (2011, chapt. 6).

Model M_h has important historical relevance to spatial capture-recapture situations (Karanth, 1995) because investigators recognized that the juxtaposition of individuals with the array of trap locations should yield heterogeneity in encounter probability, and thus it became common to use some version of Model M_h in spatial trapping arrays to estimate N. While this doesn't resolve the problem of not knowing the area relevant to N, it does yield an estimator that accommodates the heterogeneity in p induced by the spatial aspect of capture-recapture studies.

To see how this juxtaposition induces heterogeneity, we have to understand the relevance of movement in capture-recapture models. Imagine a quadrat that can be uniformly searched by a crew of biologists for some species of reptile (see Royle and Young (2008)). Figure 3.3 shows a sample quadrat searched repeatedly over a period of time. Further, suppose that species exhibits some sense of spatial fidelity in the form of a home range or territory, and individuals move about their home range (home range centroids are given by the blue dots) in some kind of random fashion. It is natural to think about it in terms of a movement process and sometimes that movement process can be modeled explicitly using hierarchical models (Royle and Young, 2008; **?**). Heuristically, we imagine that each individual in the vicinity of the study area is liable to experience variable exposure to encounter due to the overlap of its home range with the sampled area - essentially the long-run proportion of times the individual is within the sample plot boundaries, say ϕ. We might model the exposure of an individual to capture by supposing that $z_i = 1$ if individual i is available to be captured (i.e., within the survey plot) during any sample, and 0 otherwise. Then, $\Pr(z_i = 1) = \phi$. In the context of spatial studies, it is natural that ϕ should depend on *where* an individual lives, i.e., it should be individual-specific ϕ_i (Chandler et al., 2011). This system describes, precisely, that of "random temporary emigration" (Kendall, 1997) where ϕ_i is the individual-specific probability of being "available" for capture.

Conceptually, SCR models aim to deal with this problem of variable exposure to sampling due to movement in the proximity of the trapping array explicitly and formally with auxiliary spatial information. If individuals are detected with probability p_0, *conditional* on $z_i = 1$, then the marginal probability of detection

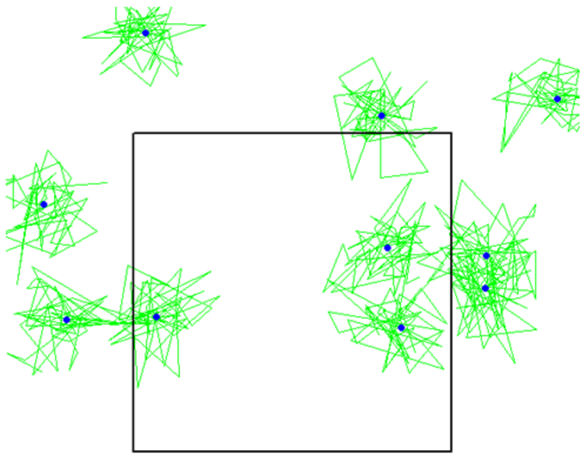

Figure 3.3: A quadrat searched for lizards and the locations of each lizard over some period of time.

of individual i is

$$p_i = p_0 \phi_i$$

so we see clearly that individual heterogeneity in encounter probability is induced as a result of the juxtaposition of individuals (i.e., their home ranges) with the sample apparatus and the movement of individuals about their home range.

We will work with a specific type of Model M_h here, that in which we extend the basic binomial observation model of Model M_0 so that

$$\text{logit}(p_i) = \mu + \eta_i$$

where

$$\eta_i \sim \text{Normal}(0, \sigma_p^2)$$

We could as well write

$$\text{logit}(p_i) \sim \text{Normal}(\mu, \sigma_p^2)$$

This "logit-normal mixture" was analyzed by Coull and Agresti (1999) and elsewhere. It is a natural extension of the basic model with constant p, as a mixed GLMM, and similar models occur throughout statistics. It is also natural to consider a beta prior distribution for p_i (Dorazio and Royle, 2003) and so-called "finite-mixture" models are also popular (Norris III and Pollock, 1996; Pledger, 2000).

3.4.1 Analysis of Model Mh

If N is known, it is worth taking note of the essential simplicity of Model M_h as a binomial GLMM. This is a type of model that is widely applied in just about every scientific discipline and using standard methods of inference based either on integrated likelihood (Laird and Ware, 1982; Berger et al., 1999) which we discuss in chapt. 6 or standard Bayesian methods. However, because N is not known, inference is somewhat more challenging. We address that here using Bayesian analysis based on data augmentation (DA). Although we use data augmentation in the context of Bayesian methods here, we note that heterogeneity models formulated under DA are easily analyzed by conventional likelihood methods as zero-inflated binomial mixtures (Royle, 2006) and more traditional analysis of model M_h based on integrated likelihood, without using data augmentation, has been considered by Coull and Agresti (1999), Dorazio and Royle (2003), and others.

As with model M_0, we have the Bernoulli model for the zero-inflation variables: $z_i \sim \text{Bern}(\psi)$ and the model of the observations expressed conditional on the latent variables z_i. For $z_i = 1$, we have a binomial model with individual-specific p_i:

$$y_i|z_i\!=\!1 \sim \text{Bin}(K, p_i)$$

and otherwise $y_i|z_i\!=\!0 \sim \delta(0)$. Further, we prescribe a distribution for p_i. Here we assume

$$\text{logit}(p_i) \sim \text{Normal}(\mu, \sigma^2)$$

The basic **BUGS** description for this model, assuming a $\text{Unif}(0, 1)$ prior for $p_0 = \text{logit}^{-1}(\mu)$, is given as follows:

```
model{

p0 ~ dunif(0,1)       # prior distributions
mup<- log(p0/(1-p0))
taup~dgamma(.1,.1)
psi~dunif(0,1)

for(i in 1:(nind+nz)){
  z[i]~dbern(psi)     # zero inflation variables
  lp[i] ~ dnorm(mup,taup) # individual effect
  logit(p[i])<-lp[i]
  mu[i]<-z[i]*p[i]
  y[i]~dbin(mu[i],J)  #  observation model
 }

N<-sum(z[1:(nind+nz)])  # N is a derived parameter
}
```

3.4.2 Analysis of the Fort Drum data

The logit-normal heterogeneity model was fitted to the bear data from the Fort Drum study, and we used data augmentation to produce a data set of $M = 300$

individuals. We ran the model using **JAGS**.

```
[... get data as before ....]

set.seed(2013)

cat("
model{
p0 ~ dunif(0,1)       # prior distributions
mup<- log(p0/(1-p0))
taup~dgamma(.1,.1)
sigmap<-sqrt(1/taup)
psi~dunif(0,1)

for(i in 1:(nind+nz)){
  z[i]~dbern(psi)     # zero inflation variables
  lp[i] ~ dnorm(mup,taup) # individual effect
  logit(p[i])<-lp[i]
  mu[i]<-z[i]*p[i]
  y[i]~dbin(mu[i],K)  #  observation model
 }

N<-sum(z[1:(nind+nz)])
}
",file="modelMh.txt")

library("rjags")
jm<- jags.model("modelMh.txt", data=data1, inits=inits, n.chains=4,
                 n.adapt=1000)
jout<- coda.samples(jm, params1, n.iter=50000, thin=1)
```

ANDY IS WORKING THIS SECTION RIGHT NOW

This produces the posterior distribution for N shown in Fig. 3.4. Posterior summaries of parameters are given as follows:

```
> summary(jout)

Iterations = 1001:201000
Thinning interval = 1
Number of chains = 4
Sample size per chain = 2e+05

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

           Mean      SD  Naive SE Time-series SE
N      106.6913 46.9809 5.253e-02       1.414777
p0       0.0867  0.0585 6.541e-05       0.001640
```

```
psi      0.2679  0.1188 1.328e-04       0.003525
sigmap   1.9279  0.4948 5.532e-04       0.014094

2. Quantiles for each variable:

             2.5%      25%      50%      75%    97.5%
N      59.000000 77.00000 93.00000 120.0000 241.0000
p0      0.005357  0.03948  0.07746   0.1244   0.2192
psi     0.140612  0.19193  0.23542   0.3043   0.6044
sigmap  1.126729  1.57266  1.86898   2.2218   3.0521
```

We used $M = 300$ for this analysis and we note that the posterior mass of N is concentrated away from this upper bound (Fig. 3.4), but there is a long right tail. Maybe or maybe not sufficient data augmentation. The model runs effectively in WinBUGS but sometimes with apparently inefficient mixing for reasons we don't understand. I did one run in WinBUGS with 200k iterations after 1k burnin and obtained this:

¿ print(fit1,digits=2) Inference for Bugs model at "modelMh.txt", fit using WinBUGS, 3 chains, each with 201000 iterations (first 1000 discarded) n.sims = 6e+05 iterations saved mean sd 2.5p0 0.09 0.06 0.01 0.04 0.08 0.13 0.22 1 96000 sigmap 1.91 0.49 1.12 1.55 1.86 2.21 3.02 1 40000 psi 0.26 0.11 0.14 0.19 0.23 0.30 0.59 1 140000 N 105.27 45.24 59.00 76.00 92.00 120.00 233.00 1 170000 deviance 184.59 17.51 153.80 172.20 183.40 195.70 222.10 1 42000

For each parameter, n.eff is a crude measure of effective sample size, and Rhat is the potential scale reduction factor (at convergence, Rhat=1).

DIC info (using the rule, pD = var(deviance)/2) pD = 153.3 and DIC = 337.8 DIC is an estimate of expected predictive error (lower deviance is better). ¿ ¿ summary(as.mcmc.list(fit1))

Iterations = 1001:201000 Thinning interval = 1 Number of chains = 3 Sample size per chain = 2e+05

1. Empirical mean and standard deviation for each variable, plus standard error of the mean:

Mean SD Naive SE Time-series SE deviance 184.59114 17.50790 2.260e-02 0.441877 N 105.27047 45.24424 5.841e-02 2.125486 p0 0.08846 0.05932 7.658e-05 0.002751 psi 0.26437 0.11456 1.479e-04 0.005298 sigmap 1.91285 0.49054 6.333e-04 0.023224

2. Quantiles for each variable:

2.5deviance 1.538e+02 172.2000 183.40000 195.7000 222.1000 N 5.900e+01 76.0000 92.00000 120.0000 233.0000 p0 5.952e-03 0.0402 0.07845 0.1277 0.2198 psi 1.399e-01 0.1900 0.23370 0.3022 0.5855 sigmap 1.118e+00 1.5520 1.86100 2.2150 3.0180

¿

The posterior mode compares well with the MLE which we obtained using the R code contained in Panel 6.1 of Royle and Dorazio (2008). The MLE of $log(n_0)$, the logarithm of the number of uncaptured individuals, is $log(n0) = 3.86$ and therefore the MLE is $\hat{N} = exp(3.86) + 47 = 94.47$ con-

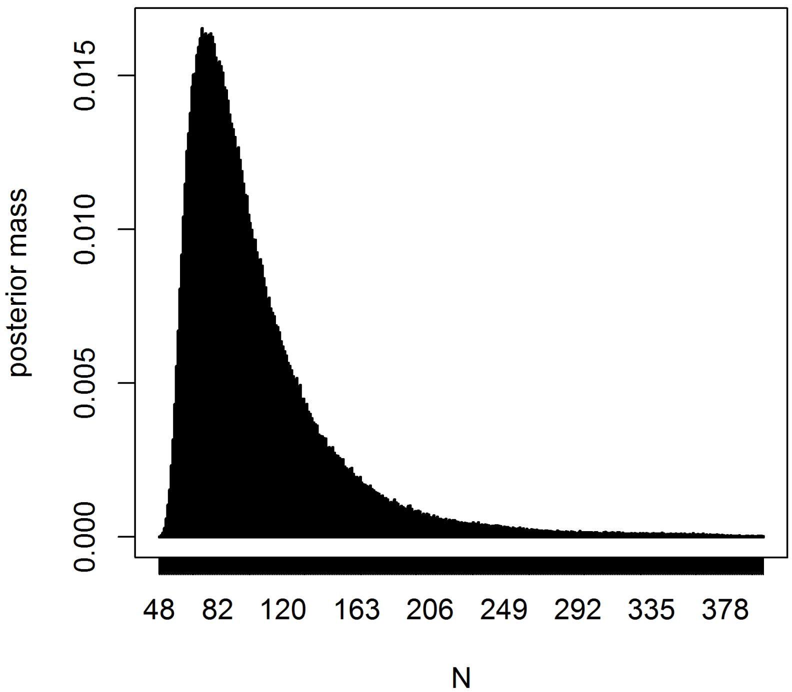

Figure 3.4: Posterior of N for Fort Drum bear study data under the logit-normal version of model M_h. From WinBUGS output. 200k samples.

sistent with the apparent mode in Figure XYZ. [5] To convert this to density we use the buffered area as computed above (255.3 km^2)[6] and perform the required summary analysis on the posterior samples of N, which results in about 0.37 individuals/km^2. The reader should carry out this analysis to confirm the estimates, and also obtain the 95% confidence interval.

3.4.3 Building your own MCMC algorithm

For fun, we construct our own MCMC algorithm using a Metropolized Gibbs sampler for Model M_h. In chapter 7 we devise MCMC algorithms for spatial

[5] We note that the result is inconsistent with Gardner et al. (2009) who reported an MLE of 104.1 ($density = 0.437 inds/km^2$) although we do not know the reason for this at the present time.

[6] WRONG #

capture-recapture models and the basic conceptual and technical considerations are entirely analogous to Model M_h.

To begin, we first collect all of our model components which are as follows: $[y_i|p_i, z_i]$, $[p_i|\mu_p, \sigma_p]$, and $[z_i|\psi]$ for *each* $i = 1, 2, \ldots, M$ and then prior distributions $[\mu_p]$, $[\sigma_p]$ and $[\psi]$.

We know the joint posterior distribution is proportional to the joint distribution of all elements y_i, p_i, z_i and also the prior distributions of the prior parameters:

$$\left\{ \prod_{i=1}^{M} [y_i|p_i, z_i][p_i|\mu_p, \sigma_p][z_i|\psi] \right\} [\mu_p, \sigma_p, \psi]$$

For prior distributions, we assume that μ_p, σ_p, ψ are mutually independent and for μ_p and σ_p we use improper uniform priors, and $\psi \sim \text{Unif}(0, 1)$. Note that the likelihood contribution for each individual, when conditioned on p_i and z_i, does not depend on ψ, μ_p, or σ_p. As such, the full-conditionals for the structural parameters ψ only depends on the collection of data augmentation variables z_i, and that for μ_p and σ_p will only depend on p_i. The full conditionals for all the unknowns are:

- (1) $[p_i|y_i, \mu_p, \sigma_p, z_i = 1] \propto [y_i|p_i][p_i|\mu_p, \sigma_p^2]$ if $z_i = 1$ $[p_i|\mu_p, \sigma_p]$ if $z_i = 0$
- (2) $z_i|\cdot \propto [y_i|z_i * p_i] * \text{Bern}(z_i|\psi)$
- (3) $\mu_p \sim \prod_i [p_i|\cdot] * \text{const}$
- (4) $\sigma_p|\cdot \sim \prod_i [p_i|\cdot] * \text{const}$
- (5) $\psi|\cdot \sim \text{Beta}(\cdot, \cdot)$

What we've done here is identify each of the full conditional distributions in sufficient detail to toss them into our Metropolis-Hastings algorithm. With the exception of ψ which has a convenient analytic solution - it is a beta distribution which we can easily sample directly. In truth, we could also sample μ_p and σ_p^2 directly with certain choices of prior distributions. For example, if $\mu_p \sim \text{Normal}(0, 1000)$ then the full conditional for μ_p is also normal.

We implement an MCMC algorithm for this model in the following block of **R** code. The basic structure is: initialize the parameters and create any required output or intermediate "holders", and then begin the main MCMC loop which, in this case, generates 100000 samples.

```
## obtain the bear data by executing the previous data grabbing
## function

temp<-getdata()
M<-temp$M
K<-temp$K
ytot<-temp$ytot
```

```
###
### MCMC algorithm for Model Mh

out<-matrix(NA,nrow=100000,ncol=4)
dimnames(out)<-list(NULL,c("mu","sigma","psi","N"))
lp<- rnorm(M,-1,1)
p<-expit(lp)
mu<- -1
p0<-exp(mu)/(1+exp(mu))
sigma<- 1
psi<- .5
z<-rbinom(M,1,psi)
z[ytot>0]<-1

for(i in 1:100000){

### update the logit(p) parameters
lpc<- rnorm(M,lp,1)  # 0.5 is a tuning parameter
pc<-expit(lpc)
lik.curr<-log(dbinom(ytot,K,z*p)*dnorm(lp,mu,sigma))
lik.cand<-log(dbinom(ytot,K,z*pc)*dnorm(lpc,mu,sigma))
kp<- runif(M) < exp(lik.cand-lik.curr)
p[kp]<-pc[kp]
lp[kp]<-lpc[kp]

p0c<- rnorm(1,p0,.05)
if(p0c>0 & p0c<1){
muc<-log(p0c/(1-p0c))
lik.curr<-sum(dnorm(lp,mu,sigma,log=TRUE))
lik.cand<-sum(dnorm(lp,muc,sigma,log=TRUE))
if(runif(1)<exp(lik.cand-lik.curr)) {
 mu<-muc
 p0<-p0c
}
}

sigmac<-rnorm(1,sigma,.5)
if(sigmac>0){
lik.curr<-sum(dnorm(lp,mu,sigma,log=TRUE))
lik.cand<-sum(dnorm(lp,mu,sigmac,log=TRUE))
if(runif(1)<exp(lik.cand-lik.curr))
 sigma<-sigmac
}
```

```
### update the z[i] variables
zc<-  ifelse(z==1,0,1)  # candidate is 0 if current = 1, etc..
lik.curr<- dbinom(ytot,K,z*p)*dbinom(z,1,psi)
lik.cand<- dbinom(ytot,K,zc*p)*dbinom(zc,1,psi)
kp<- runif(M) <  (lik.cand/lik.curr)
z[kp]<- zc[kp]

psi<-rbeta(1, sum(z) + 1, M-sum(z) + 1)

out[i,]<- c(mu,sigma,psi,sum(z))
}
```

Remarks: (1) for parameters with bounded support, i.e., σ_p and p_0, we are using a random walk candidate generator but rejecting draws outside of the parameter space. (2) We mostly use Metropolis-Hastings except for the data augmentation parameter ψ which we sample directly from its full-conditional distribution which is a beta distribution. (3) Even the latent data augmentation variables z_i are updated using Metropolis-Hastings although they too can be updated directly from their full-conditional.

3.4.4 Exercises related to model Mh

(1) Enclose the MCMC algorithm in an R function and provide arguments for some of the parameters of the function that a user might wish to modify.

(2) Execute the function and compare the results to those generated from WinBUGS in the previous section

(3) Note that the prior distribution for the "mean" parameter is given on $p_0 = exp(\mu)/(1 + exp(\mu))$. Reformulate the algorithm with a flat prior on μ and see what happens. Contemplate this.

(4) Using Bayes rule, figure out the full conditional for z_i so that you don't have to use MH for that one. It might be more efficient. Is it?

3.5 Individual Covariate Models: Toward Spatial Capture-Recapture

A standard situation in capture-recapture models is when an individual covariate is measured, and this covariate is thought to influence encounter probability. As with other closed population models, we begin with the basic binomial observation model:

$$y_i \sim \text{Bin}(K, p_i)$$

and we assume also a model for encounter probability according to:

$$\text{logit}(p_i) = \alpha_0 + \alpha_1 x_i$$

Classical examples of covariates influencing detection probability are type of animal (juvenile/adult or male/female), a continuous covariate such as body mass (Royle and Dorazio, 2008, chapt. 6), or a discrete covariate such as group or cluster size. For example, in models of aerial survey data, it is natural to model detection probabilities as a function of the observation-level individual covariate, "group size" (Royle, 2008, 2009; Langtimm, 2010).

Such "individual covariate models" are similar in structure to Model M_h, except that the individual effects are *observed* for the n individuals that appear in the sample. These models are important here because spatial capture-recapture models are precisely a form of individual covariate model, an idea that we will develop here and elsewhere. Specifically, they are such models, but where the individual covariate is a partially observed latent variable similar.. That is, unlike Model Mh, we do have some direct information about the latent variable, which comes from the spatial locations/distribution of individual recaptures. More on that later.

Traditionally, estimation of N in individual covariate models is achieved using methods based on ideas of unequal probability sampling (i.e., Horwitz-Thompson estimation), see Huggins (1989) and Alho (1990). An estimator of N is

$$\hat{N} = \sum_i \frac{1}{\tilde{p}_i}$$

where $\tilde{p}_i$ is the probability that individual i appeared in the sample. That is, $\tilde{p}_i = \Pr(y_i > 0)$. In practice, $\tilde{p}_i$ is estimated from the conditional-likelihood formed by the encounter histories. Namely,

$$\Pr(y_i | y_i > 0) = \Pr(y_i) / \Pr(y_i > 0)$$

where we substitute

$$\Pr(y_i > 0) = (1 - (1 - p_i)^K)$$

with

$$\text{logit}(p_i) = \alpha_0 + \alpha_1 x_i$$

Here we take a formal model-based approach to Bayesian analysis of such models using data augmentation (Royle, 2009). Classical likelihood analysis of the so-called "full likelihood" is covered in some detail by Borchers et al. (2002). For Bayesian analysis of individual covariate models, because the individual covariate is unobserved for the $N - n$ uncaptured individuals, we require a model to describe variation among individuals, essentially allowing the sample to be extrapolated to the population. For our present purposes, we consider a continuous covariate and we assume that it has a normal distribution:

$$x_i \sim \text{Normal}(\mu, \sigma^2)$$

Data augmentation can be applied directly to this class of models. In particular, reformulation of the model under DA yields a basic zero-inflated binomial model of the form:

$$\begin{aligned} z_i &\sim \text{Bern}(\psi) \\ y_i|z_i\!=\!1 &\sim \text{Bin}(K, p_i) \\ y_i|z_i\!=\!0 &\sim \delta(0) \end{aligned}$$

In addition, we assume that p_i is functionally related to a covariate x_i, e.g., by the logit model given above, and we assume a distribution for x_i appropriate for the context.

Fully spatial capture-recapture models essentially use this formulation with a latent covariate that is directly related to the individual detection probability (see next Section). As with the previous models, implementation is trivial in the BUGS language. The BUGS specification is very similar to that for model M_h, but we require the distribution of the covariate to be specified, along with priors for the parameters of that distribution.

3.5.1 Example: Location of capture as a covariate.

If we had a regular grid of traps over some closed geographic system then we imagine that the average location of capture would be a decent estimate (heuristically) of an individual's home range center. Intuitively some measure of typical distance from home range center to traps for an individual should be a decent covariate to explain heterogeneity in encounter probability, i.e., individuals with more exposure to traps should have higher encounter probabilities and vice versa. A version of this idea was put forth by Boulanger and McLellan (2001) (see also Ivan (2012)), but using the Huggins-Alho estimator and with covariate "distance to edge" of the trapping array. A limitation of this basic approach is that it does not provide a solution to the problem that the trap area is fundamentally ill-defined, nor does it readily accommodate the inherent and heterogeneous variation in this measured covariate. Here, we provide an example of this type of heuristically motivated approach using the fully model-based individual covariate model described above analyzed by data augmentation. We take a slightly different approach than that adopted by Boulanger and McLellan (2001). By analyzing the full likelihood and placing a prior distribution on the individual covariate, we resolve the problem of having an ill-defined area over which the population size is distributed. After you read later chapters of this book, it will be apparent that SCR models represent a formalization of this heuristic procedure.

For our purposes here, we define $x_i = ||s_i - x_0||$ where s_i is the average encounter location of individual i and x_0 is the centroid of the trap array. Conceptually, individuals in the middle of the array should have higher probability of encounter and, as x_i increases, p_i should therefore decrease. We note that we have defined s_i in terms of a sample quantity - the observed mean - which is ad hoc but maybe satisfactory under the circumstances. That said, for an expansive, dense trapping grid then we might expect the sample mean encounter location to be a good estimate of home range center but, clearly this is biased for individuals that live around the edge (or off) the trapping array. Regardless, it should be good enough for our present purposes of demonstrating this heuristically appealing application of an individual covariate model. A key point is that s_i is missing for each individual that is not encountered and thus so is x_i. Thus, it is a latent variable, or random effect, and we need therefore to specify a probability distribution for it. As a measurement of distance we know it must be positive-valued. Suppose further than we imagine no individual could have a home range radius larger than D_{max}. As such, we think a reasonable distribution for this individual covariate is

$$x_i \sim uniform(0, D_{max})$$

where D_{max} is a specified constant. In practice, people have used distance from edge of the trap array but that is less easy to define and compute.

Fort Drum Bear Study

We have to do a little bit of data processing to fit this individual covariate model to the Fort Drum data. To compute the average location of capture for each individual and the distance from the centroid of the trap array, we execute the following R instructions:

```
avg.s<-matrix(NA,nrow=nind,ncol=2)
for(i in 1:nind){
tmp<-NULL
for(j in 1:T){
aa<-bearArray[i,,j]
if(sum(aa)>0){
  aa<-  trapmat[aa>0,]
  tmp<-rbind(tmp,aa)
}
}
avg.s[i,]<-c(mean(tmp[,1]),mean(tmp[,2]))
}
Cx<-mean(trapmat[,1])
Cy<-mean(trapmat[,2])
avg.s<-rbind(avg.s,matrix(NA,nrow=nz,ncol=2))
xcent<- sqrt( (avg.s[,1]-Cx)^2 +  (avg.s[,2]-Cy)^2)
```

To define the maximum distance (maxD) from the centroid, we use that of the farthest trap, and so maxD is computed as follows:

```
minx<- min(trapmat[,1]-Cx)
maxx<-max(trapmat[,1]-Cx)
miny<- min(trapmat[,2]-Cy)
maxy<- max(trapmat[,2]-Cy)
```

```
# most extreme point determines maxD
ul<- c(minx,maxy)
maxD<- sqrt(  (ul[1]-0)^2 + (ul[2]-0)^2)
```

For the bear data the maxD was about 11.5 km. As such, the model described above will produce an estimate of the population size of bears within 11.5 units of the trap centroid[7]. The BUGS model specification and R commands to package the data and fit the model are as follows:

```
cat("
model{
p0 ~ dunif(0,1)       # prior distributions
mup<- log(p0/(1-p0))
psi~dunif(0,1)
beta~dnorm(0,.01)

for(i in 1:(nind+nz)){
  xcent[i]~dunif(0,maxD)
  z[i]~dbern(psi)     # DA variables
  lp[i] <- mup + beta*xcent[i] # individual effect
  logit(p[i])<-lp[i]
  mu[i]<-z[i]*p[i]
  y[i]~dbin(mu[i],K)  #  observation model
 }
N<-sum(z[1:(nind+nz)])
}
",file="modelMcov.txt")
data2<-list(y=ytot,nz=nz,nind=nind,K=T,xcent=xcent,maxD=11.5)
params2<-list('p0','psi','N','beta')
inits =  function() {list(z=z, psi=psi, p0=runif(1),beta=rnorm(1) ) }
fit2 = bugs(data2, inits, params2, model.file="modelMcov.txt",working.directory=getwd(
       debug=T, n.chains=3, n.iter=4000, n.burnin=1000, n.thin=4)
```

Posterior summaries are given in Table **??** XYZ, and the posterior distribution of N is given in Figure XYZ. It might be perplexing that the estimated N is much lower than obtained by model Mh but there is a good explanation for this, discussed subsequently. That issue notwithstanding, it is worth pondering how this model could be an improvement (conceptually or technically) over some other model/estimator including M0 and Mh considered previously. Well, for one, we have accounted formally for heterogeneity due to spatial location of individuals relative to exposure to the trap array, characterized by the centroid of the array. Moreover, we have done so using a model that is based on an explicit mechanism, as opposed to a phenomenological one such as Model Mh. Moreover, importantly, using our new model, *the estimated N applies to an explicit area which is defined by our prescribed value of maxD.* That is, this area

[7] To be convincing this might need a little bit of hand-holding

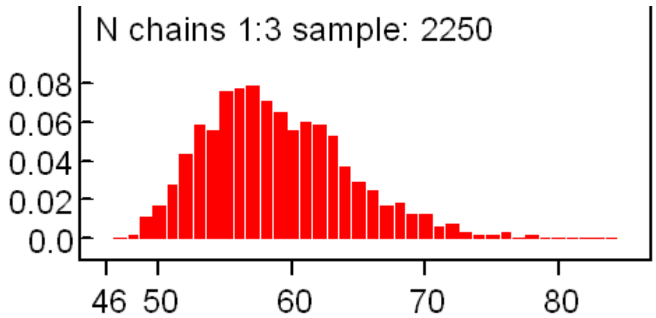

Figure 3.5: Needs a caption

is a fixed component of the model and the parameter N therefore has explicit spatial context, as the number of individuals with home range centers less than maxD from the centroid of the trap array. As such, the implied “effective trap area”[8] for any maxD is that of a circle with radius maxD.

```
%% Not sure whether this should be a table or verbatim print-out
\begin{table}
\tabular{ccccccccc}
Node statistics
node  mean  sd  MC error 2.5% median 97.5% start sample
N 58.89 5.483 0.2199 50.0 58.0 71.0 251 2250
beta -0.246 0.06087 0.003892 -0.3592 -0.2457 -0.126 251 2250
deviance 459.4 13.29 0.4496 435.7 458.4 487.8 251 2250
p0 0.5409 0.06817 0.004052 0.4072 0.544 0.6678 251 2250
psi 0.1706 0.02572 7.759E-4 0.1247 0.1692 0.2242 251 2250
\end{tabular}
\caption{..... xyz ......}
\end{table}
\label{tab.maxD}
```

We’ll remake this figure in R. For now, insert it as is.

3.5.2 Extension of the Model

One important issue in understanding the meaning of estimates produced under the individual covariate model is that the uniform distribution on maxD implies that density is *not constant* over space. In particular, this model implies that it *decreases* as we move away from the centroid of the trap array. This is one reason we have a lower estimate of density than that obtained previously and also why, if we were to increase maxD, we would see density continue to decrease: $x[i] \sim$

[8]This is a bad use of this term. We have never defined ETA or ESA. What is it, exactly?

Uniform$(0, maxD)$ implies constant N in each distance band from the centroid but obviously the *area* of each distance band is increasing. The reader can verify this as a homework exercise. Obviously, the use of an individual covariate model is *not* restricted to use of this specific distribution for the individual covariate. Clearly, it is a bad choice and, therefore, we should think about whether we can choose a better distribution for maxD - one that doesn't imply a decreasing density as distance from the centroid increases. Conceptually, what we want to do is impose a prior on distance from the centroid, x, such that density is proportional to the amount of area in each successive distance band as you move farther away from the centroid. In fact, there is theory that exists which tells us what the correct distribution of x is $2x/maxD^2$. This can be derived by noting that $F(x) = Pr(X < x) = pi * x * x/pi * maxD * maxD$. Then, $f(x) = dF/dx = 2 * x/(maxD * maxD)$. This might be called a triangular distribution, I think, which makes sense because the incremental area in each additional distance band increases linearly with radius (i.e., distance from centroid). It is sometimes comforting to verify things empirically:

```
> u<-runif(10000,-1,1)
> v<-runif(10000,-1,1)
> d<- sqrt(u*u+v*v)
> hist(d[d<1])
> hist(d[d<1],100)
> hist(d[d<1],100,probability=TRUE)
> abline(0,2)
```

It would be useful if we could describe this distribution in *BUGS but there is not a built-in way to do this. One possibility is to use a discrete version of the pdf. We might also be able to use what is referred to in WinBUGS jargon as the "zeros trick" (see Advanced BUGS tricks) although we haven't pursued this approach. Instead, we consider using a discrete version and break Dmax into L distance classes of width δ, with probabilities proportional to $2 * x$. In particular, if the cut-points are $xg[1] = 0, xg[2], \ldots, xg[L+1] = Dmax$ and the interval midpoints are $xm[i] = xg[i+1] - \delta$. Then, the interval probabilities are $p[i] = 2 * xm[i] * delta/(Dmax * Dmax)$, which we can compute once and then send them to WinBUGS as data.

The R script is as follows. In the model description the variable x (observed home range center) has been rounded so that the discrete version of the $f(x)$ can be used as described previously. The new variable labeled `xround` is actually then the integer category label in units of delta from 0. Thus, to convert back to distance in the expression for $lp[i]$, `xround[i]` has to be multiplied by δ.

```
delta<-.2
xround<-xcent%/%delta  + 1
Dgrid<- seq(delta,maxD,delta)
xprobs<- delta*(2*Dgrid/(maxD*maxD))
xprobs<-xprobs/sum(xprobs)
```

Table 3.4: Table: Analysis of Fort Drum bear hair snare data using the individual covariate model, for different values of Dmax, the upper limit of the uniform distribution of 'distance from centroid of the trap array'

maxD mn SD [1,] 12 0.230 0.038 [2,] 15 0.244 0.041 [3,] 17 0.249 0.044 [4,] 18 0.249 0.043 [5,] 19 0.250 0.043 [6,] 20

```
cat("
model{
p0 ~ dunif(0,1)       # prior distributions
mup<- log(p0/(1-p0))
psi~dunif(0,1)
beta~dnorm(0,.01)

for(i in 1:(nind+nz)){
  xround[i]~dcat(xprobs[])
  z[i]~dbern(psi)                     # zero inflation variables
  lp[i] <- mup + beta*xround[i]*delta # individual effect
  logit(p[i])<-lp[i]
  mu[i]<-z[i]*p[i]
  y[i]~dbin(mu[i],K)  #  observation model
 }

N<-sum(z[1:(nind+nz)])
}
",file="modelMcov.txt")
```

To fit the model we do this - keeping in mind that the data objects required below have been defined in previous analyses of this chapter:

```
data2<-list(y=ytot,nz=nz,nind=nind,K=T,xround=xround,xprobs=xprobs,delta=delta)
params2<-list('p0','psi','N','beta')
inits =  function() {list(z=z, psi=psi, p0=runif(1),beta=rnorm(1) ) }
fit = bugs(data2, inits, params2, model.file="modelMcov.txt",working.directory=getwd(),
       debug=FALSE, n.chains=3, n.iter=11000, n.burnin=1000, n.thin=2)
```

This is a useful model because it induces a clear definition of area in which the population of N individuals reside. Under this model, that area is defined by specification of maxD. We can apply the model for different values of maxD and observe that the estimated N varies with $maxD$. Fortunately, we see empirically, that while N seems highly sensitive to the prescribed value of $maxD$, density seems to be invariant to $maxD$ as long as it is chosen to be sufficiently large. We fit the model for $maxD = 12$ (points in close proximity to the trap arra) to 20 for and the results are given in Table **??**.

We see that the posterior mean and SD of density (individuals per square km) appear insensitive to choice of $maxD$ once we get a slight ways away from the maximum observed value of about 11.5. The estimated density of 0.250

per km^2 is actually quite a bit lower than we reported using model Mh (0.37, see section XYZ above) for which sample area is not an explicit feature of the model. On the other hand it is higher than that reported from Model M0 using the buffered area (0.195). There is no basis really for comparing or contrasting these various estimates and it would be a useful philosophical exercise for the reader to discuss this matter. In particular, application of model M0 and Mh are distinctly *not* spatially explicit models – the area within which the population[9] resides is not defined under either model. There is therefore no reason at all to think that the estimates produced under either model, using a buffered area, are justifiable based on any theory. In fact, we would get exactly the same estimate of N no matter what we declare the area to be. On the other hand, the individual covariate model explicitly describes a distribution for "distance from centroid" that is a reasonable and standard null model - it posits, in the absence of direct information, that individual home range centers are randomly distributed in space and that probability of detection depends on the distance between home range center and the centroid of the trap array. Under this definition of the system, we see that density is invariant to the choice of sample area which seems like a desirable feature. The individual covariate model is not ideal, however, because it does not make full use of the spatial information in the data set, i.e., the trap locations and the locations of each individual encounter.

3.5.3 Invariance of density to maxD

Under the model above, and also under models that we consider in later chapters, a general property of the estimators is that while N increases with the prescribed trap area (equivalent to maxD in this case), we expect that density estimators should be invariant to this area. In the model used above, we note that $Area(maxD) = \pi * maxD * maxD$ and $E[N(maxD)] = \lambda * A(maxD)$ and thus $E[Density(maxD)] = \lambda$ which is constant. This should be interpreted as the *prior* density. Absent data, then realizations under the model will have density λ regardless of what $maxD$ is prescribed to be. As we verified empirically above, the posterior density is also invariant if $maxD$ as long as the implied area (implied by maxD) is large enough so that the data no longer provide information about density (i.e., "far away"), then our estimator of density should become insensitive.

3.5.4 Toward Fully Spatial Capture-recapture Models

We developed this model for the average observed location and equated it to home range center s_i. Intuitively, taking the average encounter location as an estimate of home range center makes sense but more so when the trapping grid is dense and expansive relative to typical home range sizes. However, our approach also ignored the variable precision with which each $s[i]$ is estimated and also, as noted previously, estimates of $s[i]$ around the "edge" (however we define that)

[9] We need to look back at Chapter 1 and make sure we quit calling this "sample area" - it really isn't that at al, but rather the area within which N resides.

are biased because the observations are truncated (we can only observe locations within the trap array). In the next Chapter we provide a further extension of this individual covariate model that definitively resolves the ad hoc nature of the individual covariate approach we took here. In that model we build a model in which $s[i]$ are regarded as latent variables and the observation locations (i.e., trap specific encounters) are linked to those latent variables with an explicit model. We note that the model fitted previously could be adapted easily to deal with s_i as a latent variable, simply by adding a prior distribution for s_i. The reader should contemplate how to do this in WinBUGS.

3.6 DISTANCE SAMPLING: A primative Spatial Capture-Recapture Model

Distance sampling is one of the most popular methods for estimating animal abundance. One of the great benefits of distance sampling is that it provides explicit estimates of *density*. The distance sampling model is a special case of a closed population model with a covariate. The covariate in this case, x_i, is the distance between an individual's location "u" and the observation location or transect. In fact, the model underlying distance sampling is precisely the same model as that which applies to the individual-covariate models, except that observations are made at only $K = 1$ sampling occasion. In a sense, distance sampling is a spatial capture-recapture model, but without the "recapture." This first and most basic spatial capture-recapture model has been used routinely for decades and, formally, it is a spatially-explicit model in the sense that it describes, explicitly, the spatial organization of individual locations (although this is not always stated explicitly) and, as a result, somewhat general models of how individuals are distributed in space can be specified (Royle, 2004; Johnson, 2010; Sillett, 2011). As before, the distance sampling model, under data augmentation, includes a set of M zero-inflation variables z_i and the binomial model expressed conditional on z (binomial for $z = 1$, and fixed zeros for $z = 0$). In distance sampling we pay for having only a single sample (i.e., $K = 1$) by requiring constraints on the model of detection probability. A standard model is

$$\log(p_i) = b * x_i^2$$

for $b < 0$, where x_i denotes the distance at which the ith individual is detected relative to some reference location where perfect detectability ($p = 1$) is assumed. This function corresponds to the "half-normal" detection function (i.e., with $b = 1/\sigma^2$). If $K > 1$ then the intercept alpha is identifiable and such models are usually called "capture-recapture distance sampling"(Borchers, missing) and others XYZ????).

As with previous examples, we require a distribution for the individual covariate x_i. The customary choice is

$$x_i \sim \text{Uniform}(0, B)$$

wherein $B > 0$ is a known constant, being the upper limit of data recording by the observer (i.e., the point count radius, or transect half-width). In practice, this is sometimes asserted to be infinity, but in such cases the distance data are usually truncated. Specification of this distance sampling model in the BUGS language is shown in Panel 3.1. Royle and Dorazio (2008), p. xyz) provide a distance sampling example analyzed by DA using the famous Impala data.

```
b~dunif(0,10)
psi~dunif(0,1)

for(i in 1:(nind+nz)){
   z[i]~dbern(psi)    # DA Variables
   x[i]~dunif(0,B)    # B=strip width
   p[i]<-exp(logp[i])   # DETECTION MODEL
   logp[i]<-   -((x[i]*x[i])*b)
   mu[i]<-z[i]*p[i]
   y[i]~dbern(mu[i])  # OBSERVATION MODEL
 }
N<-sum(z[1:(nind+nz)])
D<- N/striparea  # area of transects
```

Panel 3.1: Distance sampling model in WinBUGS, using a “half-normal” detection function.

As with the individual covariate model in the previous section, the distance sampling model can be equivalently specified by putting a prior distribution on individual *location* instead of distance between individual and observation point (or transect). Thus we can write the general distance sampling model as

$$logit(p[i]) = alpha + beta * ||u[i] - x0||$$

Along with

$$\mathbf{u}_i \sim Uniform(\mathcal{S})$$

where x_0 is a fixed point (or line) and $u[i]$ is the individual's location which is observable for n individuals. In practice it is easier to record distance instead of location. Basic math can be used to argue that if individuals have a uniform distribution in space, then the distribution of Euclidean distance is also uniform. In particular, if a transect of length L is used and x is distance to the transect then $F(x) = Pr(X <= x) = L * x/L * B = x/B$ and $f(x) = dF/dx = (1/B)$. For measurements of radial distance, see the previous section.

In the context of our general characterization of SCR models (chapter 1.XYZ), we suggested that every SCR model can be described, conceptually, by a hierarchical model of the form:

$$[y|u][u|s][s].$$

Distance sampling ignores s, and treats u as observed data[10]. Thus, we are left with

$$[y|u][u].$$

In contrast, as we will see in the next chapters, basic SCR models (chapter 4) ignore u and condition on s, which is not observed:

$$[y|s][s]$$

Since $[u]$ and $[s]$ are both assumed to be uniformly distributed, these are structurally equivalent models! The main differences have to do with interpretation of model components and whether or not the latent variables are observable (in distance sampling they are).

So why bother with SCR models when distance sampling yields density estimates and accounts for spatial heterogeneity in detection? For one, imagine try to collect distance sampling data on tigers! Clearly, distance sampling requires that one can collect large quantities of distance data, which is not always possible. For tigers, it is much easier, efficient, and safer to employ camera traps or tracking plates and then apply SCR models. Furthermore, as we will see in Ch XYZ, SCR models can use distance data to estimate all the parameters of our enchilada, allowing us to study distribution, movement, and density. Thus, SCR models are much more flexible than distance sampling models, and can accommodate data from virtually all animal survey designs.

3.6.1 Example: Muntjac deer survey from Nagarahole, India

Here we fit distance sampling models to distance sampling data on the muntjac deer (Muntiakus muntjak) collected in the year 2004 from Nagarahole National Park in southern India (Kumar, missing)(Kumar et al. unpublished data). The muntjac is a solitary species and distance measurements were made on 57 groups that were largely singletons with XYZ pairs of individuals. Commands for reading in and organizing the data for WinBUGS, followed by writing the model to a text file. Note that the total sampled area of the transects is fed in as "striparea" which is 708 (km of transect) multiplied by the strip width (B=150 = 0.15 km) multiplied by 2.

```
library("R2WinBUGS")
data<- read.csv("Muntjac.csv")
nind<-nrow(data)
y<-rep(1,nind)
nz<-400
y<-c(y,rep(0,nz))
x<-data[,3]
x<-c(x,rep(NA,nz))
z<-y
```

[10] Formally we could also say that $[u] = \int [y|s][s]ds$

```
data<-list(y=y,x=x,nz=nz,nind=nind,B=150,striparea=708*.15*2)

cat("
model{
b~dunif(0,10)
psi~dunif(0,1)

for(i in 1:(nind+nz)){
   z[i]~dbern(psi)    # DA Variables
   x[i]~dunif(0,B)    # B=strip width
   p[i]<-exp(logp[i])   # DETECTION MODEL
   logp[i]<-   -((x[i]*x[i])*b)
   #logp[i]<- -b*log(x[i]+1)
   mu[i]<-z[i]*p[i]
   y[i]~dbern(mu[i])  # OBSERVATION MODEL
 }
N<-sum(z[1:(nind+nz)])
D<- N/striparea  # area of transects
}
",file="dsamp.txt")
```

Next, we provide inits, indicate which parameters to monitor, and then pass those things to WinBUGS:

```
params<-list('b','N','D','psi')
inits =  function() {list(z=z, psi=runif(1), b=runif(1,0,.02) )}
fit = bugs(data, inits,  params, model.file="dsamp.txt",
working.directory=getwd(),debug=T, n.chains=3, n.iter=4000, n.burnin=1000, n.thin=2)
```

Posterior summaries are provided in the following table. Estimated density is pretty low, 1.1 individuals per sq. km.[11]

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
D	1.096	0.1694	0.009122	0.8098	1.078	1.474	501	4500
N	232.8	35.99	1.938	172.0	229.0	313.0	501	4500
b	5.678E-4	1.05E-4	4.129E-6	3.867E-4	5.616E-4	7.949E-4	501	4500
deviance	681.2	16.72	0.7536	650.8	680.6	716.6	501	4500
psi	0.5099	0.08238	0.004442	0.3681	0.5033	0.6918	501	4500

3.7 Summary and Outlook

Traditional closed population capture-recapture models are closely related to binomial generalized linear models. Indeed, the only real distinction is that in capture-recapture models, the population size parameter N (corresponding also

[11] much lower than Samba's : Observers walked about 708 km from 39 transects in Nagarahole and the muntjac density is about 3 per sq km.. I need to get to the bottom of this.

to the size of a hypothetical "complete" data set) is unknown. This requires special consideration in the analysis of capture-recapture models. The classical approach to inference recognizes that the observations don't have a standard binomial distribution but, rather, a truncated binomial (from which which the so-called "conditional likelihood" derives) since we only have encounter frequency data on observed individuals. If instead we analyze the models using data augmentation, the observations can be modeled using a zero-inflated binomial distribution. In short, when we deal with the unknown-N problem using data augmentation then we are left with zero-inflated GLM and GLMMs instead of ordinary GLM or GLMMs. The analysis of such zero-inflated models is practically convenient, especially using the various Bayesian analysis packages that use the BUGS language.

Spatial capture-recapture models that we will consider in the rest of the chapters of this book are closely related to what have been called individual covariate models. Heuristically, spatial capture-recapture models arise by defining individual covariates based on observed locations of individuals – we can think of using some function of mean encounter location as an individual covariate. We did this in a novel way, by using distance to the centroid of the trapping array as a covariate. We analyzed the "full likelihood" using data augmentation, and placed a prior distribution on the individual covariate which was derived from an assumption that individual locations are, a priori, uniformly distributed in space. This assumption provides for invariance of the density estimator to the choice of population size area (induced by maximum distance from the centroid of the). The model addressed some important problems in the use of closed population models: it allows for heterogeneity in encounter probability due to the spatial context of the problem and it also provides a direct estimate of density because area is a feature of the model (via the prior on the individual covariate). The model is still not completely general because the model does not make use of the fully spatial encounter histories, which provide direct information about the locations and density of individuals. A specific individual covariate model that is in widespread use is classical "distance sampling." The model underlying distance sampling is precisely a special kind of SCR model - but one without replicate samples. Understanding distance sampling and individual covariate models more broadly provides a solid basis for understanding and analyzing spatial capture-recapture models.

Chapter 4

Fully Spatial Capture-Recapture Models

Chapter 5

Other observation models

Chapter 6

Maximum likelihood estimation

Chapter 7

MCMC details

Chapter 8

Goodness of Fit and stuff

Chapter 9

Covariate models

Chapter 10

Inhomogeneous Point Process

Chapter 11

Open models

Bibliography

Alho, J. (1990), "Logistic regression in capture-recapture models," *Biometrics*, 623–635.

Bales, S. L., Hellgren, E. C., Leslie Jr., D. M., and Hemphill Jr., J. (2005), "Dynamics of recolonizing populations of black bears in the Ouachita Mountains of Oklahoma," *Wildlife Society Bulletin*, 1342–1351.

Berger, J. O., Liseo, B., and Wolpert, R. L. (1999), "Integrated likelihood methods for eliminating nuisance parameters," *Statistical Science*, 1–22.

Borchers, D. L. (missing), "missing," *Missing*, missing.

Borchers, D. L., Buckland, S. T., and Zucchini, W. (2002), *Estimating animal abundance: closed populations*, vol. 13, Springer Verlag.

Boulanger, J. and McLellan, B. (2001), "Closure violation in DNA-based mark-recapture estimation of grizzly bear populations," *Canadian Journal of Zoology*, 79, 642–651.

Burnham, K. P. and Overton, W. S. (1978), "Estimation of the size of a closed population when capture probabilities vary among animals," *Biometrika*, 65, 625.

Chandler, R., Royle, J., and King, D. (2011), "Inference about density and temporary emigration in unmarked populations," *Ecology*, 92, 1429–1435.

Converse, S. J. and Royle, J. A. (2010), "Missing," *Missing*, Missing.

Coull, B. A. and Agresti, A. (1999), "The Use of Mixed Logit Models to Reflect Heterogeneity in Capture-Recapture Studies," *Biometrics*, 55, 294–301.

Dice, L. R. (1938), "Some census methods for mammals," *Journal of Wildlife Management*, 2, 119–130.

Dorazio, R. and Royle, J. (2003), "Mixture models for estimating the size of a closed population when capture rates vary among individuals," *Biometrics*, 351–364.

Durbin and Elston (2012), "Missing," *Missing*, Missing.

Fienberg, S. E., Johnson, M. S., and Junker, B. W. (1999), "Classical multilevel and Bayesian approaches to population size estimation using multiple lists," *Journal of the Royal Statistical Society of London A*, 163, 383–405.

Gardner, B. (2009), "missing," *Missing*, missing.

Gardner, B., Royle, J., Wegan, M., Rainbolt, R., and Curtis, P. (2010), "Estimating black bear density using DNA data from hair snares," *The Journal of Wildlife Management*, 74, 318–325.

Gopalaswamy (2012), "Missing," *Missing*, missing.

Huggins, R. M. (1989), "On the statistical analysis of capture experiments," *Biometrika*, 76, 133.

Illian (2008), "Missing," *Missing*, Missing.

Ivan, J. (2012), "Density, demography, and seasonal movements of snowshoe hares in central Colorado," Ph.D. thesis, COLORADO STATE UNIVERSITY.

Johnson (2010), "Missing," *missing*, missing.

Karanth, K. U. (1995), "Estimating tiger Panthera tigris populations from camera-trap data using capture–recapture models," *Biological Conservation*, 71, 333–338.

Kendall (1997), "Missing," *Missing*, missing.

Kéry, M., Gardner, B., Stoeckle, T., Weber, D., and Royle, J. A. (2010), "Use of Spatial Capture-Recapture Modeling and DNA Data to Estimate Densities of Elusive Animals," *Conservation Biology*, 25, 356–364.

Kery, M. and Schaub, M. (2011), *Bayesian Population Analysis Using WinBugs*, Academic Press.

King, R. (missing), "Missing," *missing*, Missing.

Kumar (missing), "Unpublished data," .

Laird, N. M. and Ware, J. H. (1982), "Random-effects models for longitudinal data," *Biometrics*, 963–974.

Langtimm (2010), "Missing," *Missing*, missing.

Link, W. A. (2003), "Missing," *missing*, missing.

Liu and Wu (1999), "Parameter expansion for data augmentation," *J Am Stat Assoc*, 94, 1264–1274.

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J. A., and Langtimm, C. A. (2002), "Estimating site occupancy rates when detection probabilities are less than one," *Ecology*, 83, 2248–2255.

Mackenzie, D. I. and Royle, J. (2005), "Designing occupancy studies: general advice and allocating survey effort," *Journal of Applied Ecology*, 42, 1105–1114.

Norris III, J. L. and Pollock, K. H. (1996), "Nonparametric MLE under two closed capture-recapture models with heterogeneity," *Biometrics*, 639–649.

Pledger, S. (2000), "Unified maximum likelihood estimates for closed capture-recapture models using mixtures," *Biometrics*, 434–442.

Royle, J. (2006), "Site occupancy models with heterogeneous detection probabilities," *Biometrics*, 62, 97–102.

— (2009), "Analysis of capture-recapture models with individual covariates using data augmentation," *Biometrics*, 65, 267–274.

Royle, J. and Dorazio, R. (2008), *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities*, Academic Press.

— (2010), "Missing," *Missing*, missing, missing.

— (2011), "Missing," *Missing*, missing, missing.

Royle, J., Dorazio, R., and Link, W. (2007), "Analysis of multinomial models with unknown index using data augmentation," *Journal of Computational and Graphical Statistics*, 16, 67–85.

Royle, J. A. (2004), "Missing," *Missing*, missing.

— (2008), "Modeling individual effects in the Cormack–Jolly–Seber model: a state–space formulation," *Biometrics*, 64, 364–370.

— (2010), "missing," *missing*, missing.

Royle, J. A. and Young, K. V. (2008), "A Hierarchical Model For Spatial Capture-Recapture Data," *Ecology*, 89, 2281–2289.

Sanathanan, L. (1972), "Estimating the size of a multinomial population," *The Annals of Mathematical Statistics*, 142–152.

Schofield and Barker (missing), "Missing," *Missing*, missing.

Sillett (2011), "Missing," *Missing*, missing.

Tanner, M. A. and Wong, W. H. (1987), "The calculation of posterior distributions by data augmentation," *J Am Stat Assoc*, 82, 528–540.

Tyre, A. J., Tenhumberg, B., Field, S. A., Niejalke, D., Parris, K., and Possingham, H. P. (2003), "Improving precision and reducing bias in biological surveys: estimating false-negative error rates," *Ecological Applications*, 13, 1790–1801.

Wegan (missing), “missing,” *missing*, missing.

Yang, H. C. and Chao, A. (2005), “Modeling Animals’ Behavioral Response by Markov Chain Models for Capture–Recapture Experiments,” *Biometrics*, 61, 1010–1017.