

4. OCCUPANCY AND ABUNDANCE (在・不在データと個体数)

岡村 寛

背景

- occupancy (presence/absence) dataは効率的で重要であるが、不正確で粗い

local population = high abundanceならば、個体群全体が大きく減少しても、occupancy dataはなかなか感知できない

- $\psi = \Pr(z=1)$: 個体数 N に関係. 個体数が大きければ、そのareaがoccupyされている確率は高い

背景

- N: local population size
 $g(N|\theta)$: probability mass function for N
- A: area size
 $A \rightarrow 0$ then $g(N) \rightarrow 0$ or 1 i.e. $N(A) \in \{0,1\}$
この場合, occupancy = abundance
 $g(N)$: Poisson $\rightarrow \psi = 1 - \exp(-\lambda)$
 $\lambda \rightarrow 0 \Rightarrow \psi \approx \lambda$ Aはコントロール可なところが重要

Fig. 4.1.

```
par(las=1)
```

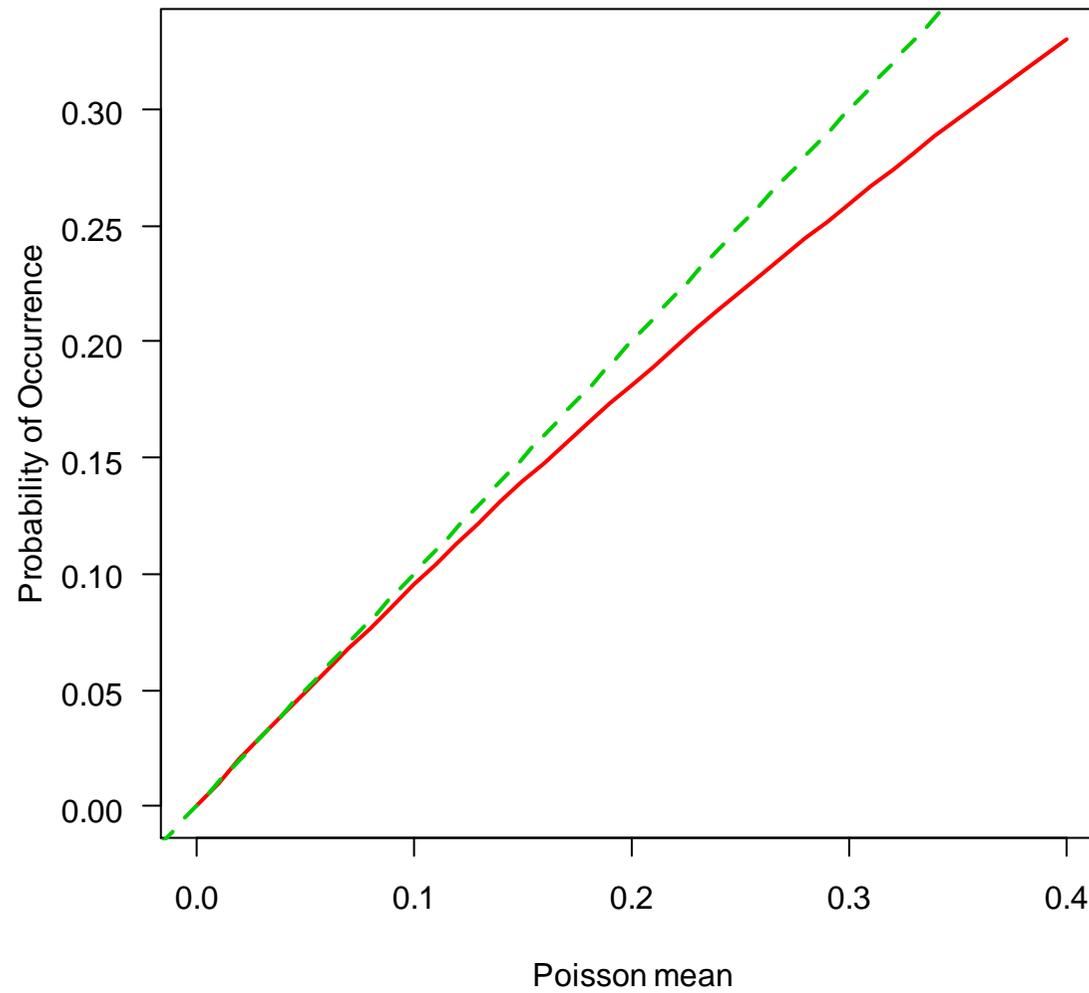
```
psi <- function(lambda) 1-exp(-lambda)
```

```
x <- seq(0,0.4,by=0.01)
```

```
plot(x, psi(x), xlab="Poisson mean", ylab="Probability of  
Occurrence",col=2,type="l",lwd=2)
```

```
abline(a=0,b=1,lty=2,col=3,lwd=2)
```

Fig. 4.1.



背景

- occupancyとabundanceの階層モデル

$\Pr(N)=g(N|\theta)$: abundanceのモデル

$\Pr(y|N)=f(y|N, r)$: presence/absenceの観測モデル

Contents

- 4.1 Abundance-induced heterogeneity in detection
- 4.2 Prediction of local abundance
- 4.3 Modeling covariate effects
- 4.4 Other model extensions
- 4.5 Functional independence between P and Abundance
- 4.6 The variable area sampling design
- 4.7 Estimating occupancy in the absence of replicate samples
- 4.8 Summary

4.1

- Chap. 3のoccupancy modelの発見確率 p は一定か既知の共変量の関数だった
- ここでは, p が未知の量(個体数)の関数である場合を考える.
- site間の p の変化はsite間のabundanceの変化の反映
- 在/不在データからlocal abundanceを推定する

4.1.1. RN model

- site: $i = 1, 2, \dots, M$
- replication for each site: $j = 1, 2, \dots, J$
- response variable: $y_{ij} \in \{0, 1\}$
- N_i : population size at site i , **latent variable**
- **site間でabundanceが異なるのでsite間で発見確率が異なる(abund大 \rightarrow detect prob大)**

4.1.1. RN model

- r : 1個体が発見される確率
- N_i : site i のabundanceとするとき, site i の発見確率 $p_i = 1 - (1 - r)^{N_i}$
 $1 - r$ = 発見なしの確率
 $(1 - r)^{N_i}$ は, N_i 全部発見されない
 $1 - (1 - r)^{N_i}$ んなこたない(少なくとも1匹は発見)

4.1.1. RN model

- $y_i \sim \text{Bi}(J, p_i) = J! / \{y_i!(J-y_i)!\} p(N_i)^{y_i} (1-p(N_i))^{J-y_i}$
- N に何も仮定をおかないとふつつう推定不可能
- $N_i \sim g(N | \theta)$
- $g(N | \theta) = \text{Po}(\lambda)$
- $\lambda_i = \lambda$ or $\log(\lambda_i) = \beta_0 + \beta_1 x_i$
- x_i は時間とか
- ポアソンモデルはecologyのカウントモデルで最もよくある選択

4.1.2. Derived parameters

- モデル内に直接 ψ は出てこない
- $\psi = \Pr(N > 0) = 1 - g(0 | \theta)$
- $p_i = p(N_i)$: $N_i = 0 \rightarrow p_i = 0$
- 一方, Chap. 3の p は, $p = \Pr(y=1 | z=1)$. つまり, $N > 0$ のとき.

4.1.2. Derived parameters

$$\begin{aligned} p_{rn} &= \Pr(y > 0) \\ &= \Pr(y > 0 | N=0)\Pr(N=0) + \Pr(y > 0 | N > 0)\Pr(N > 0) \\ &= \Pr(y > 0 | N > 0)\Pr(N > 0) \\ &= p_c \psi \end{aligned}$$

Chap. 3 $\mathcal{D}p$

4.1.3. Induced heterogeneity

- $N_i = 0 \Rightarrow p_i = 0$
 - $N_i = 1 \Rightarrow p_i = r$
 - $N_i = 2 \Rightarrow p_i = 1 - (1 - r)^2$
 - ...
 - $N_i = \infty \Rightarrow p_i = 1$
-
- Model “ M_h ” (再捕確率に個体差) と類似
“ M_h ” では $p \sim$ ベータ分布, $\text{logit}(p) \sim$ 正規分布を使う
が, RN model では $p = p(N)$, $N \sim g(N|\theta)$

4.1.3. Induced heterogeneity

- N がabundanceであると解釈しなければならない理由はない

(4.1.5)

- $p = 1 - (1 - r)^N$
- $\log(1 - p) = N \log(1 - r)$
- $\log(-\log(1 - p)) = \log(-\log(1 - r)) + \log(N)$

4.1.4. Analysis by integrated likelihood

- 推定の際, ランダム効果 N_i は積分して消される
 - $q(y|r, \theta)$
 - = $\sum_{N=0}^{\infty} f(y|p(N))g(N|p)$
 - = $\sum_{N=1}^{\infty} f(y|p(N))g(N|p)$
 - + $I(y=0)g(0|\theta)$ (**zero-inflated form, pp.111-112**)
- g: Poisson $\Rightarrow q(y|r, \theta) = \sum f(y|N, r)e^{-\lambda}\lambda^N/N!$

4.1.5. Bird Point Counts

- BBS: North American Breeding Bird Survey
- roadside survey at 50 locations ($M = 50$)
- $J = 11$
- response: 11回中のcatbirdの発見数



4.1.5. Bird Point Counts

```
> source("catbird.y")
```

```
> table(factor(catbird.y,levels=0:11))
```

```
0 1 2 3 4 5 6 7 8 9 10 11
```

```
31 5 5 5 2 0 1 1 0 0 0 0
```

1回のvisitで2個体以上発見されたのは550回中6回だけだった。よって、 $N(A) \in \{0, 1\}$ に近く、在・不在データにより失われた情報は小さい

Panel 4.1

```
expit <- function(x) 1/(1+exp(-x))

RN.fn <- function(y,J=11,M=50,Nmax=100){
####
lik<-function(parms){
  r<-expit(parms[1]) # 個体の発見確率
  lambda<-exp(parms[2]) # ポアソン平均
  pvec<-1-(1-r)^(0:Nmax) # p
  gN<-dpois(0:Nmax,lambda) # ポアソン 0~Nmax
  gN<-gN/sum(gN) # ポアソン確率修正
```

Panel 4.1

```
lik<-rep(NA,M) # 尤度
for(i in 1:M){
  lik[i]<-sum(dbinom(y[i],J,pvec)*gN)
}
-1*sum(log(lik))
}
```

Panel 4.1

```
####
```

```
tmp<-nlm(lik,c(0,0),hessian=TRUE) # 最適化  
ests<-tmp$estimate  
aic<-tmp$minimum*2 + 2*length(ests)  
se<- sqrt(diag(solve(tmp$hessian)))  
list(ests=ests,se=se,aic=aic)  
}
```

4.1.5. Bird Point Counts

```
> catbird.y
```

```
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
0 0 2 7 1 6 3 0 1 3 0 2 0 4 1 0 0 0 0 0 0 0 0 0
25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48
0 0 0 0 3 0 0 1 4 2 3 0 0 0 0 0 2 2 1 0 0 0 3 0
49 50
0 0
```

4.1.5. Bird Point Counts

```
> res4.1 <- RN.fn(catbird.y)
```

```
> res4.1
```

```
$ests
```

```
[1] -1.521006 -0.605054
```

```
$se
```

```
[1] 0.2396590 0.2504918
```

```
$aic
```

```
[1] 135.9431
```

4.1.5. Bird Point Counts

```
> r.hat <- expit(res4.1$ests[1])
```

```
> r.hat
```

```
[1] 0.1793134
```

```
> lambda.hat <- exp(res4.1$ests[2])
```

```
> lambda.hat
```

```
[1] 0.5460449
```


4.1.5. Bird Point Counts

siteごとの最大カウントの平均は0.46

```
> (sum(catbird.y>0)+4)/50
```

```
[1] 0.46
```

(2個体以上のカウントがあったのは6回なので、2つのsiteで2回2個体の観測が、また1つのsiteで3回2個体の観測があった)

```
> lambda.hat/0.46
```

```
[1] 1.187054
```

4.1.5. Bird Point Counts

カウントデータがない場合:

```
> sum(catbird.y>0)
```

```
[1] 19
```

$\Pr(y > 0) = 1 - \exp(-\lambda)$ より,

```
> -log(1-19/50)
```

```
[1] 0.4780358
```

4.1.5. Bird Point Counts

two-state model

```
lik<-function(parms){  
  ...  
  pvec<-1-(1-p)^(0:1)  
  pvec<-c(pvec[1],rep(pvec[2],Ninfty))  
  ...  
  for(i in 1:M){  
    lik[i]<-sum(dbinom(y[i],J,pvec)*gN)  
  }  
  ...  
}
```

4.1.5. Bird Point Counts

```
> exp(-0.6632238)
```

```
[1] 0.5151878 # lambda for two-state model
```

```
0.546 > 0.515
```

Why?: p に変動があるとき, p を一定とすると, p は過大推定される(個体数がでかいのほど発見されやすいから. Chap.6, p.193)

```
> expit(-1.2065374)
```

```
[1] 0.2303143 # p for two-state model
```

4.1.5. Bird Point Counts

```
> prn <- sum((1-(1-  
r.hat)^(0:100))*dpois(0:100,lambda.hat))
```

```
> prn
```

```
[1] 0.09326504
```

```
> psi <- 1-dpois(0,lambda.hat)
```

```
> psi
```

```
[1] 0.4207638
```

```
> pc <- prn/psi
```

```
> pc
```

```
[1] 0.2216565
```

4.1.6. Application to carnivore survey data

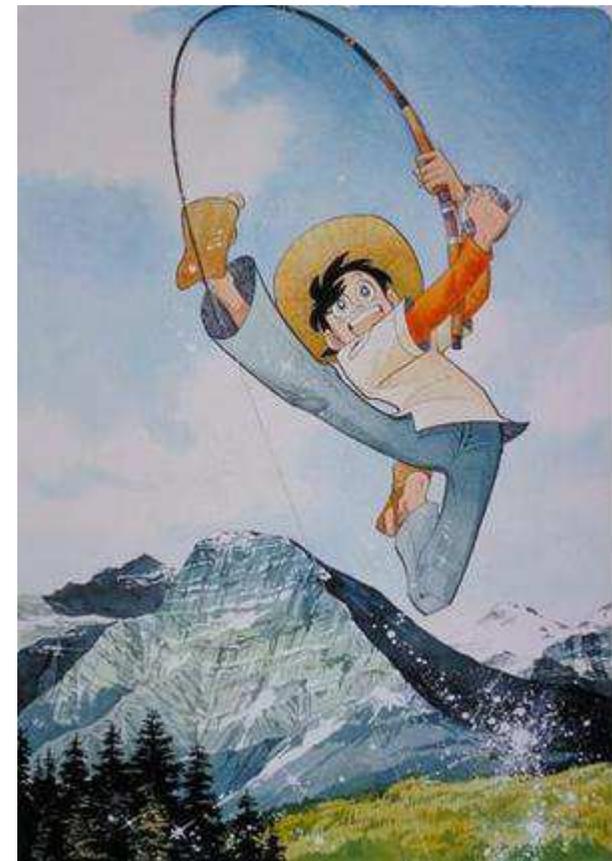
- Forest carnivore survey data

M = 464 stations

J = 8

data

(400,16,12,12,5,10,3,4,2)



4.1.6. Application to carnivore survey data

```
fn.4.1.6.1 <-  
function(y=0:8,wt=c(400,16,12,12,5,10,3,4,2),J=8,Nmax=100){  
lik<-function(parms){  
  p<-expit(parms[1]); lambda<-exp(parms[2])  
  pvec<-1-(1-p)^(0:1)  
  pvec<-c(pvec[1],rep(pvec[2],Nmax))  
  gN<-dpois(0:Nmax,lambda); gN<-gN/sum(gN)  
  lik<-rep(NA,J+1)  
  for(i in 1:(J+1)){  
    lik[i]<-sum(dbinom(y[i],J,pvec)*gN)  
  }  
-1*sum(wt*log(lik))  
}  
nlm(lik,c(0,0),hessian=TRUE)  
}
```

4.1.6. Application to carnivore survey data

```
> res4.1.6.1 <- fn.4.1.6.1()
> expit(res4.1.6.1$est[1])
[1] 0.3993689 # p
> 1-dpois(0,exp(res4.1.6.1$est[2]))
[1] 0.1403072 # psi
> sum(y[-1])/464
[1] 0.137931 # apparent occupancy
> 1-(1-0.3993689)^8
[1] 0.983062 # power
```


4.1.6. Application to carnivore survey data

```
fn.4.1.6.2 <-  
function(y=0:8,wt=c(400,16,12,12,5,10,3,4,2),J=8,Nmax=100){  
lik<-function(parms){  
  r<-expit(parms[1]); lambda<-exp(parms[2])  
  pvec<-1-(1-r)^(0:Nmax)  
  gN<-dpois(0:Nmax,lambda); gN<-gN/sum(gN)  
  lik<-rep(NA,J+1)  
  for(i in 1:(J+1)){  
    lik[i]<-sum(dbinom(y[i],J,pvec)*gN)  
  }  
  -1*sum(wt*log(lik))  
}  
nlm(lik,c(-0.5,-1.8),hessian=TRUE)  
}
```

4.1.6. Application to carnivore survey data

```
> res4.1.6.2 <- fn.4.1.6.2()
```

```
> expit(res4.1.6.2$est[1])
```

```
[1] 0.3682166 # r
```

```
> exp(res4.1.6.2$est[2])
```

```
[1] 0.1621897 # lambda
```

```
> 1-dpois(0, 0.1621897)
```

```
[1] 0.1497201 # psi
```

4.1.6. Application to carnivore survey data

```
> prn <- sum((1-(1-  
0.3682166)^(0:100))*dpois(0:100,0.1621897))
```

```
> prn
```

```
[1] 0.05797262
```

```
> psi <- 1-dpois(0,0.1621897)
```

```
> psi
```

```
[1] 0.1497201
```

```
> pc <- prn/psi
```

```
> pc
```

```
[1] 0.3872066
```

4.1.6. Application to carnivore survey data

Ninftyの選び方に注意が必要だが, Ninftyを指定する必要がない方法もある.

Monien, H. (2006) Gaussian summation: an exponentially convergent summation scheme. arXiv:math/0611057v1. Available at http://arxiv.org/PS_cache/math/pdf/0611/0611057v1.pdf.

Rプログラム:

<https://sites.google.com/site/hiroshiokamura/home>

4.2

- random effects そのものに興味がある

⇒ Best Unbiased Prediction (BUP)

By Bayes' rule,

$$[N | y, r, \lambda] = f(y | N, r)g(N | \lambda) / q(y | r, \lambda)$$

$$N^{\{\hat{\forall}\}} = \forall \sum k \Pr(N=k | y, r, \lambda)$$

r, λ の不確実性はbootstrapなどで考慮可

Table 4.2.

```
bup <- function(r, lambda, J=11, kmax=100){
  pvec <- 1 - (1-r)^(0:kmax)
  gN <- dpois(0:kmax, lambda)
  gN <- gN/sum(gN)
  out <- matrix(NA, ncol=kmax+1, nrow=J+1)
  for(i in 0:kmax){
    for(j in 0:J){
      out[j+1, i+1] <- dbinom(j, J, pvec[i+1]) * gN[i+1]
    }
  }
  norm.c <- rowSums(out)
  apply(out, 2, function(x) x/norm.c)
}
```

Table 4.2.

```
> table4.2 <- bup(r=0.179, lambda=0.546)
> Table4.2 <-
round(cbind(table4.2[,1:9],rowSums(table4.2[,9:
100])),3)
> colnames(Table4.2) <- c(0:8,"9+")
> rownames(Table4.2) <- 0:11
# E[N | y]
> apply(table4.2,1,function(x) sum(x*0:100))
```

4.2.1. What is N?

- N = abundance per unit area
- N = effective population size
- N の解釈問題があり, GLMなどでそれを回避することは可能だが, N をbiologicallyに解釈する方が好ましいと考える

4.3.

$$N_i \sim \text{Pois}(\lambda_i)$$

$$\log(\lambda_i) \sim \beta_0 + \beta_1 x_i$$

$$p_{ij} = 1 - (1 - r_{ij})^{N_i}$$

$$\text{logit}(r_{ij}) = \alpha_0 + \alpha_1 x_{ij}$$

$$\mathbf{y}_i = (1, 1, 0)$$

$$\Pr(\mathbf{y}_i | \alpha_0, \alpha_1, N_i) = p_{i1} p_{i2} (1 - p_{i3})$$

$$\Pr(\mathbf{y}_i | \alpha_0, \alpha_1, \lambda) = \sum_{N_i=0}^{\infty} p_{i1} p_{i2} (1 - p_{i3}) \Pr(N_i | \lambda)$$

4.3.1. Analysis of the Swiss Willow Tit Data

```
RNlikcov<-function(parms,vars=c("a0","b0")) ){
  ones<-rep(1,M)
  tmp<-c(0,0,0,0,0,0,0,0,0)
  nam<-
c("b0","a0","intensity","date1","date2","elev1","elev2","forest","length")
  names(tmp)<-nam
  tmp[vars]<-parms
  b0<-tmp[1];a0<-tmp[2];a1<-tmp[3];a2<-tmp[4]; a3<-tmp[5];
  b1<-tmp[6];b2<-tmp[7];b3<-tmp[8];b4<-tmp[9]

  rvec<-expit(a0*ones + a1*intensity + a2*date + a3*date2)
  lamvec<-exp(b0*ones + b1*elev + b2*elev2 + b3*forest + b4*length)
```

4.3.1. Analysis of the Swiss Willow Tit Data

```
lik<-rep(NA,M)
for(i in 1:M){
  gN<-dpois(0:Nmax,lamvec[i])
  gN<-gN/sum(gN)
  dvec<-ymat[i,]
  naflag<-is.na(dvec)
  PMAT<- 1-outer( (1-rvec[i,]),0:Nmax,"^")
  LIK<- t((PMAT^dvec)*((1-PMAT)^(1-dvec)))
  LIK[,naflag]<-1
  likvec<-apply(LIK,1,prod)
  lik[i]<-sum(likvec*gN)
}
-1*sum(log(lik))
}
```

4.3.1. Analysis of the Swiss Willow Tit Data

```
> panel4pt3.fn() # Table 4.3
```

```
Null model
```

```
$minimum
```

```
[1] 260.9785
```

```
$estimate
```

```
[1] 1.0804977 -0.8558065
```

```
...
```

```
Null model + intensity + DATE1 + DATE2 + LENGTH + FOREST + ELEV + ELEV2
```

```
$minimum
```

```
[1] 198.1841
```

```
$estimate
```

```
[1] 0.23689770 -0.47141124 0.48970494 -0.08305434 0.01182408
```

```
[6] 0.27829552 0.58843663 1.91118220 -0.94703238
```

4.4.

- Model extensions
non-binomial relationship between p and N

4.1. Non-binomial detection

$$\Pr(y=1 | N_i) = 1 - (1 - r)^{\theta N_i}$$

θ は非独立性を表すパラメータ

$$\text{cloglog}(p) = \text{cloglog}(r) + \log(\theta) + \log(N_i)$$

$\text{cloglog}(r)$ と $\log(\theta)$ を区別できない

区別できる形 (quasi-binomial detection model)

$$\Pr(y=1 | N_i) = 1 - (1 - r)^{N_i^\theta}$$

Panel 4.4.

```
lik<-function(parms){
  r<-expit(parms[1])
  lambda<-exp(parms[2])
  alpha<-exp(parms[3])
  pvec<-1- (1-r)^( (0:Nmax)^(1+alpha) ) #  $\theta = 1+\alpha$ 
  gN<-dpois(0:Nmax,lambda)
  gN<-gN/sum(gN)
  lik<-rep(NA,nsites)
  for(i in 1:nsites){
    lik[i]<-sum(dbinom(y[i],J,pvec)*gN)
  }
  -1*sum(log(lik))
}
```

4.4.2. Alternative abundance models

- $g(N | \alpha, \mu) =$

$$\frac{\Gamma(N+\alpha)}{\{\Gamma(\alpha)N!\}} \left\{ \frac{\alpha}{\alpha+\mu} \right\}^\alpha \left\{ \frac{\mu}{\alpha+\mu} \right\}^N$$

$$E(N) = \mu, \text{Var}(N) = \mu + \mu^2/\alpha, \text{disp. par. } \varepsilon = 1/\alpha$$

問題

- ・非現実的な仮定: 個体数が増加すると分散/平均が増加
- ・収束不安定

4.4.2. Alternative abundance models

- Poisson/log-normal mixture

$$N \sim \text{Po}(\lambda), \log(\lambda) \sim N(\mu, \sigma)$$

$$q(N | r, \mu, \sigma) = \sum_{N=0}^{\infty} p(N)$$

$$f(y | p(N)) = \int_{\phi} g(N | \phi) h(\phi | \mu, \sigma) d\phi$$

$$\phi = \log(\lambda)$$

Panel 4.5.

```
lik<-function(parms){
  r<-expit(parms[1]); mu<-parms[2]; sigma<-exp(parms[3])
  uy<-unique(y)
  il<-rep(NA,Nmax+1)
  for(k in 0:Nmax){
    il[k+1]<-integrate(
      function(x){
        dpois(k,exp(x))*dnorm(x,mu,sigma)
      }, lower=-Inf,upper=Inf)$value
  }
  il2<-rep(NA,J+1)
  for(k in uy){
    il2[k+1]<- sum(dbinom(k,J,1-(1-r)^(0:Nmax))*il)
  }
  -1*sum(log(il2[y+1] ))
}
```

4.4.3. Example

- Catbird data

Table 4.4

N Model	detection	r or p	$E(N)$	ψ	extra parm	AIC
Po	bin	0.179	0.546	0.421	–	135.94
Po	qb	0.164	0.547	0.421	$\alpha = 0.29$	137.57
Po/log-Norm	bin	0.153	0.655	0.426	$\sigma = 0.72$	137.67
NB	bin	0.149	0.674	0.426	$\epsilon = 0.67$	137.68
Two-state	const	0.230	0.515	0.403	–	138.52

4.5.

- p が N と独立な場合
- two-state model (Chap.3)に帰する
- $g(N|\theta)$ を推定できない場合, 補助情報が必要 (variable area)
- variable area情報を使う場合, replicationが必要なくなる

4.5.1. Abundance and occupancy under independence of p and N

- $\Pr(y | p, N) = \sum \Pr(y | p) \Pr(N)$
 $= \sum_{N=1} \Pr(y | p) \Pr(N) + I(y=0) \Pr(N=0)$
 $= \Pr(y | p) \sum_{N=1} \Pr(N) + I(y=0) \Pr(N=0)$
 $= \Pr(y | p) \Pr(N > 0) + I(y=0) \Pr(N=0)$

(Chap. 3の式(3.5.1))

g : Poisson $\Rightarrow \psi(\lambda) = 1 - \exp(-\lambda)$

$$q(y | p, \lambda) = f(y | p) \psi(\lambda) + I(y=0)(1 - \psi(\lambda))$$

λ は推定可能だが, g の形の区別はできない

個体数の分布は分からない

4.5.2. Choice of link functions

- 必ずしもlogit-linkだけではない

$g(N|\theta) = (1-\theta)\theta^N$: 幾何分布

$$\psi = 1 - \Pr(N=0) = \theta, E(N) = \mu = \theta/(1 - \theta)$$

$$\log(\mu) = \log(\theta/(1 - \theta)) = \bar{\text{logit}}(\psi)$$

$g(N|\lambda)$: Poisson

$$\psi = 1 - \exp(-\lambda)$$

$$\text{cloglog}(\psi) = \log(\lambda)$$

4.5.3. Assessing functional independence

- p と N が独立のとき, RN model \rightarrow two-state model

しかし, 2つのモデル間に包含関係はない \Rightarrow AICの使用

4.6.

- p と N が独立の時, N の分布は分からない
しかし, 異なる(しかし既知の)area情報が使え
れば, 分布も知ることが出来る

$$N_i \sim NB(\mu A_i, \alpha)$$

$$q(y_i | p, \alpha, \mu, A_i)$$

$$= f(y_i | J, p) \psi(\alpha, \mu | A_i) + I(y=0) \{1 - \psi(\alpha, \mu | A_i)\}$$

$$\psi(\alpha, \mu | A) = 1 - g(0 | \alpha, \mu, A) = 1 - (\alpha / (\alpha + A\mu))^\alpha$$

4.6.1. Illustration

- シミュレーションでうまくいくか検討 (Table 4.5)

Table 4.6

case	λ_0	p	α	'Power' against alternative				median($\hat{\lambda}_0$)		
				Poi/ind	NB/dep	Poi/dep	NB/ind	Poi/ind	NB/dep	Poi/dep
1	0.1	0.4	1.0	0.28	0.905	0.805	0.096	0.078	0.079	0.080
2	0.1	0.6	1.0	0.32	0.925	1.000	0.098	0.078	0.076	0.076
3	0.2	0.4	1.0	0.65	0.990	1.000	0.205	0.132	0.135	0.135
4	0.2	0.6	1.0	0.68	0.985	1.000	0.200	0.131	0.125	0.125
5	0.1	0.4	0.5	0.59	0.910	0.140	0.116	0.064	0.065	0.067
6	0.1	0.6	0.5	0.62	0.925	0.570	0.113	0.065	0.064	0.064
7	0.2	0.4	0.5	0.87	0.970	0.990	0.217	0.103	0.104	0.106
8	0.2	0.6	0.5	0.88	0.980	1.000	0.216	0.102	0.098	0.099

4.7.

N: Poisson

$$P(A) = \Pr(y=1 | A) = \Pr(\text{発見かつ } N > 0) = p(1 - e^{-A\lambda})$$

N: NB

$$P(A) = p\{1 - (\alpha / (\alpha + A\lambda))^\alpha\}$$

$A \rightarrow \infty$ のとき, $p(A) \rightarrow p$

y_1, \dots, y_M があれば, $y \sim \text{Bi}(1, p(A))$

PoissonとNBの区別をするためには数百以上の観測値が必要

4.8. Summary

- 在不在データとabundanceのlink ~階層モデル
- p と N が独立のとき, RN model \rightarrow two-state model. しかし, N の分布を推定できない.
- Variable area情報があれば, two-state modelで N の分布も推定できる. さらに, replicationも必要ない.
- model-based推定の重要性