

Modelos de ocorrência: uma espécie, várias estações

Contents

Preparação	1
Ajuste dos modelos	2
Seleção de modelos	8
Cálculo do previsto	8

-
- [Arquivo em pdf](#)
 - [Arquivo em markdown](#) (para executar os comandos no R studio)
-

Preparação

Abra o R e carregue os pacotes necessários

```
library(unmarked)
library(RMark)
library(stringr)
library(plyr)
```

Usaremos dados de registros da coruja manchada do norte *Strix occidentalis caurina* em 55 sítios de floresta temperada nos EUA, por 5 anos. [Aqui](#) há mais informações sobre este caso de estudo.

Os dados estão no formato nativo do MARK (.inp). Use os comandos abaixo para convertê-los para um objeto da classe *unmarkedMultFrame*, do pacote *unmarked*:

```
## Link dos dados na página da disciplina
url <- "http://ecologia.ib.usp.br/bie5703/lib/exe/fetch.php?media=roteiros:corujas.inp"
## Importa arquivo inp
tmp <- convert.inp(url)
## Seleciona historico de capturas e converte em data frame
y <- str_split(tmp$ch, pattern="")
y <- ldply(y, as.numeric)[,2:41]
```

Os dados são de oito visitas anuais, por 5 anos. Temos então que criar covariável que indique o ano de cada visita, que deve ser uma matriz com número de linhas igual ao número de sítios (55 no caso) e número de linhas igual ao número de anos.¹

```
ano <- matrix(data = rep(factor(1998:2002), nrow(y)),
              nrow=nrow(y), ncol=ncol(y)/8, byrow=TRUE)
```

¹neste exemplo a única covariável é a identidade do ano em que ocorreu cada visita. Se há outras covariáveis associadas aos anos basta criar outras matrizes com os valores destas covariáveis com as mesmas dimensões.

Como neste exemplo não há covariáveis dos sítios nem das visitas já podemos criar o objeto para ajustar o modelo ². Note que o número de ocasiões (anos no caso) deve ser informado no argumento `numPrimary`:

```
## Cria objeto para o modelo de ocupação multiseason do unmarked
## (Veja vinhetas para detalhes)
coruj <- unmarkedMultFrame(y = y,
                           yearlySiteCovs = list(ano=ano),
                           numPrimary=5)
```

```
## Verifica objeto
summary(coruj)
```

```
## unmarkedFrame Object
##
## 55 sites
## Maximum number of observations per site: 40
## Mean number of observations per site: 26.33
## Number of primary survey periods: 5
## Number of secondary survey periods: 8
## Sites with at least one detection: 46
##
## Tabulation of y observations:
##      0      1 <NA>
## 1052  396  752
##
## Yearly-site-level covariates:
##      ano
## 1998:55
## 1999:55
## 2000:55
## 2001:55
## 2002:55
```

Ajuste dos modelos

O pacote *unmarked* usa a [sintaxe de modelos lineares](#) do R e tem funções para diferentes tipos de modelos de ocupação. Consulte as vinhetas do pacote para mais informações

```
## Lista da vinhetas
vignette(package="unmarked")
## Abre pdf da vinheta de introdução
vignette(topic="unmarked", package="unmarked")
## Abre pdf da vinheta dos modelos de múltiplas estações
vignette(topic="colext", package="unmarked")
```

Para os modelos de ocupação com covariáveis usamos função *colext*. Seus primeiros quatro argumentos são fórmulas para os parâmetros de ocupação inicial (ψ), colonização (γ), extinção (ϵ) e detecção (p).

O modelos mais simples é o que tem todos estes parâmetros constantes:

²Veja `help(colext)` e `vignette(topic="colext", package="unmarked")` para incluir covariáveis de sítios e de visitas.

```
## ~1 indica constante
coruj.m1 <- colext(psiformula=~1, gammaformula=~1,
                  epsilonformula=~1, pformula=~1,
                  data=coruj)
## Resumo do modelo
summary(coruj.m1)
```

```
##
## Call:
## colext(psiformula = ~1, gammaformula = ~1, epsilonformula = ~1,
##       pformula = ~1, data = coruj)
##
## Initial (logit-scale):
## Estimate SE z P(>|z|)
## 0.537 0.289 1.86 0.063
##
## Colonization (logit-scale):
## Estimate SE z P(>|z|)
## -1.49 0.284 -5.23 1.65e-07
##
## Extinction (logit-scale):
## Estimate SE z P(>|z|)
## -1.73 0.26 -6.66 2.73e-11
##
## Detection (logit-scale):
## Estimate SE z P(>|z|)
## -0.021 0.0743 -0.283 0.777
##
## AIC: 1363.32
## Number of sites: 55
## optim convergence code: 0
## optim iterations: 23
## Bootstrap iterations: 0
```

```
## Coeficientes na escala logito
coef(coruj.m1)
```

```
## psi(Int) col(Int) ext(Int) p(Int)
## 0.53719570 -1.48831083 -1.72876244 -0.02103079
```

```
## Intervalos de confiança dos coeficientes
confint(coruj.m1, type='det') #p
```

```
## 0.025 0.975
## p(Int) -0.1667289 0.1246674
```

```
confint(coruj.m1, type="psi") #psi inicial
```

```
## 0.025 0.975
## psi(Int) -0.02911498 1.103506
```

```
confint(coruj.m1, type="col") #gamma
```

```
##           0.025      0.975  
## col(Int) -2.045534 -0.9310872
```

```
confint(coruj.m1, type="ext") #epsilon
```

```
##           0.025      0.975  
## ext(Int) -2.237498 -1.220027
```

O padrão dos modelos de ocupação é usar a função logito para as probabilidades de detecção e ocupação:

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right)$$

Portanto os coeficientes retornados pelas funções `summary` e `coef` estão nesta escala. Para obter as probabilidades estimadas pelo modelo na escala original use a função `backTransform`.

```
## Coeficientes na escala de probabilidades  
(m1.p <- backTransform(coruj.m1, type="det"))
```

```
## Backtransformed linear combination(s) of Detection estimate(s)  
##  
## Estimate      SE LinComb (Intercept)  
##    0.495 0.0186  -0.021          1  
##  
## Transformation: logistic
```

```
(m1.psi= backTransform(coruj.m1, type="psi"))
```

```
## Backtransformed linear combination(s) of Initial estimate(s)  
##  
## Estimate      SE LinComb (Intercept)  
##    0.631 0.0673   0.537          1  
##  
## Transformation: logistic
```

```
(m1.gamma = backTransform(coruj.m1, type="col"))
```

```
## Backtransformed linear combination(s) of Colonization estimate(s)  
##  
## Estimate      SE LinComb (Intercept)  
##    0.184 0.0427  -1.49          1  
##  
## Transformation: logistic
```

```
(m1.epsilon = backTransform(coruj.m1, type="ext"))
```

```
## Backtransformed linear combination(s) of Extinction estimate(s)
##
## Estimate      SE LinComb (Intercept)
##      0.151 0.0332   -1.73           1
##
## Transformation: logistic
```

```
## Intervalos de confiança
confint(m1.p)
```

```
##      0.025      0.975
## 0.4584141 0.5311265
```

```
confint(m1.psi)
```

```
##      0.025      0.975
## 0.4927218 0.7509165
```

```
confint(m1.gamma)
```

```
##      0.025      0.975
## 0.1145044 0.2827042
```

```
confint(m1.gamma)
```

```
##      0.025      0.975
## 0.1145044 0.2827042
```

Vamos ajustar um modelo em que detecção e colonização variam entre ocasiões (anos)³:

```
coruj.m2 <- colext(psiformula=~1, gammaformula=~ano-1,
                  epsilonformula=~1, pformula=~ano-1,
                  data=coruj)
## Resumo do modelo
summary(coruj.m2)
```

```
##
## Call:
## colext(psiformula = ~1, gammaformula = ~ano - 1, epsilonformula = ~1,
##       pformula = ~ano - 1, data = coruj)
##
## Initial (logit-scale):
## Estimate SE      z P(>|z|)
##      0.53 0.286 1.85 0.0638
##
## Colonization (logit-scale):
##      Estimate SE      z P(>|z|)
## ano1998  -2.109 0.768 -2.75 0.00602
```

³A notação ano-1 elimina o intercepto da covariável, que corresponde ao primeiro ano. É apenas uma conveniência para ter um coeficiente para cada ano.

```
## ano1999    -2.541 0.984 -2.58 0.00978
## ano2000    -0.453 0.452 -1.00 0.31590
## ano2001    -2.284 1.050 -2.17 0.02963
##
## Extinction (logit-scale):
## Estimate    SE      z P(>|z|)
##      -1.82 0.272 -6.69 2.19e-11
##
## Detection (logit-scale):
##      Estimate    SE      z P(>|z|)
## ano1998    0.3604 0.164  2.203 0.0276
## ano1999    0.0824 0.161  0.511 0.6093
## ano2000   -0.3364 0.180 -1.865 0.0622
## ano2001   -0.5243 0.184 -2.848 0.0044
## ano2002    0.1479 0.157  0.941 0.3469
##
## AIC: 1352.105
## Number of sites: 55
## optim convergence code: 0
## optim iterations: 38
## Bootstrap iterations: 0
```

E agora com a extinção diferente entre ocasiões e colonização constante:

```
coruj.m3 <- colext(psiformula=~1, gammaformula=~1,
                  epsilonformula=~ano-1, pformula=~ano-1,
                  data=coruj)
## Resumo do modelo
summary(coruj.m3)
```

```
##
## Call:
## colext(psiformula = ~1, gammaformula = ~1, epsilonformula = ~ano -
##      1, pformula = ~ano - 1, data = coruj)
##
## Initial (logit-scale):
## Estimate    SE      z P(>|z|)
##      0.508 0.285 1.78 0.0746
##
## Colonization (logit-scale):
## Estimate    SE      z P(>|z|)
##      -1.51 0.29 -5.21 1.88e-07
##
## Extinction (logit-scale):
##      Estimate    SE      z P(>|z|)
## ano1998    -2.28 0.625 -3.66 0.000256
## ano1999    -1.92 0.593 -3.23 0.001224
## ano2000    -1.15 0.459 -2.50 0.012381
## ano2001    -1.99 0.592 -3.37 0.000763
##
## Detection (logit-scale):
##      Estimate    SE      z P(>|z|)
## ano1998    0.3678 0.163  2.259 0.0239
```

```
## ano1999  0.0838 0.162  0.516  0.6060
## ano2000 -0.3846 0.187 -2.055  0.0399
## ano2001 -0.4317 0.173 -2.496  0.0126
## ano2002  0.1433 0.158  0.905  0.3656
##
## AIC: 1357.025
## Number of sites: 55
## optim convergence code: 0
## optim iterations: 47
## Bootstrap iterations: 0
```

E com extinção e colonização diferentes entre anos:

```
coruj.m4 <- colext(psiformula=~1, gammaformula=~ano-1,
                  epsilonformula=~ano-1, pformula=~ano-1,
                  data=coruj)
## Resumo do modelo
summary(coruj.m4)
```

```
##
## Call:
## colext(psiformula = ~1, gammaformula = ~ano - 1, epsilonformula = ~ano -
##      1, pformula = ~ano - 1, data = coruj)
##
## Initial (logit-scale):
##      Estimate      SE      z P(>|z|)
##      0.53 0.286 1.86 0.0635
##
## Colonization (logit-scale):
##      Estimate      SE      z P(>|z|)
## ano1998 -2.116 0.770 -2.75 0.00598
## ano1999 -2.599 0.997 -2.61 0.00916
## ano2000 -0.463 0.447 -1.04 0.29981
## ano2001 -2.028 0.843 -2.40 0.01619
##
## Extinction (logit-scale):
##      Estimate      SE      z P(>|z|)
## ano1998 -2.33 0.623 -3.74 0.000183
## ano1999 -1.87 0.579 -3.22 0.001277
## ano2000 -1.16 0.477 -2.42 0.015343
## ano2001 -2.00 0.593 -3.38 0.000722
##
## Detection (logit-scale):
##      Estimate      SE      z P(>|z|)
## ano1998  0.3613 0.163  2.211 0.02703
## ano1999  0.0773 0.161  0.479 0.63175
## ano2000 -0.3450 0.184 -1.875 0.06075
## ano2001 -0.4715 0.183 -2.576 0.00999
## ano2002  0.1445 0.158  0.913 0.36121
##
## AIC: 1355.65
## Number of sites: 55
## optim convergence code: 0
```

```
## optim iterations: 48
## Bootstrap iterations: 0
```

Seleção de modelos

O *unmarked* tem funções para criar uma lista de modelos e então realizar sua seleção por diversos critérios

```
modelos <- fitList("p(.)psi(.)e(.)g(.)"=coruj.m1,
                  "p(ano)psi(.)e(.)g(ano)"=coruj.m2,
                  "p(ano)psi(.)e(ano)g(.)"=coruj.m3,
                  "p(ano)psi(.)e(ano)g(ano)"=coruj.m4)
modSel(modelos)
```

```
##              nPars      AIC delta  AICwt cumltvWt
## p(ano)psi(.)e(.)g(ano)      11 1352.10  0.00 0.7943      0.79
## p(ano)psi(.)e(ano)g(ano)     14 1355.65  3.55 0.1349      0.93
## p(ano)psi(.)e(ano)g(.)      11 1357.03  4.92 0.0678      1.00
## p(.)psi(.)e(.)g(.)          4 1363.32 11.22 0.0029      1.00
```

O modelo de menor AIC (e portanto $\Delta AIC = 0$) é o mais plausível. Convenciona-se que modelos com $\Delta AIC \leq 2$ são tão plausíveis quanto o selecionado.

Cálculo do previsto

Valor previsto dos parâmetros

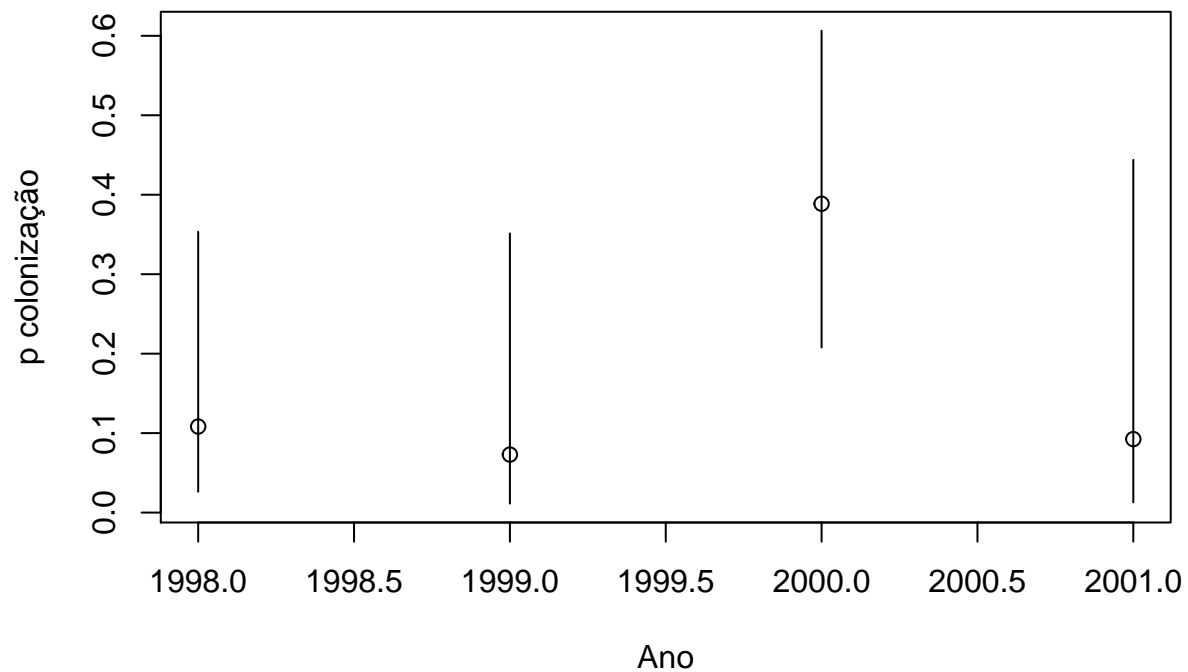
O modelo com variação de γ entre anos é o único com $\Delta AIC \leq 2$. Usamos a função `pred` para estimar os valores de colonização ao longo dos anos:

```
## primeiro criamos um dataframe com os valores das covariaveis em que faremos as previsões
## Objeto com as covariaveis
df1 <- data.frame(ano=factor(1998:2001))
## Previstos e seus Se e ICs
(coruj.m4.pred <- predict(coruj.m2, type='col', newdata = df1))
```

```
##      Predicted      SE      lower      upper
## 1 0.10824860 0.07410853 0.02625079 0.3534163
## 2 0.07299968 0.06657395 0.01132116 0.3513065
## 3 0.38861130 0.10735206 0.20771917 0.6064519
## 4 0.09246675 0.08811843 0.01284361 0.4437936
```

E um exemplo de gráfico dos previstos e seus intervalos de confiança:

```
## Juntando os previstos as covariaveis
coruj.m4.pred$ano <- 1998:2001
## Plot de psi x distância para o habitat Creosote
plot(Predicted ~ ano, data=coruj.m4.pred,
     ylim=range(coruj.m4.pred[,3:4]), ylab="p colonização", xlab="Ano")
with(coruj.m4.pred, segments(x0=ano, y0=lower, x1=ano, y1=upper))
```

Valor previsto de parâmetros derivados

Nos modelos de ocupação com múltiplas ocasiões estima a probabilidade inicial de ocupação ψ_1 . A probabilidade de ocupação na ocasião seguinte ψ_2 é:

$$\psi_2 = \psi_1\phi_1 + (1 - \phi_1)\gamma$$

Em que ϕ é a probabilidade de persistência ($\phi_i = 1 - \epsilon_i$).

Portanto, as probabilidades de ocupação para as ocasiões exceto a primeira são **parâmetros derivados**. O objeto resultante do ajuste já tem estas quantidades guardadas nele. Para vê-las digite

```
projected(coruj.m2)
```

```
##              1              2              3              4              5
## unoccupied 0.3704589 0.4181376 0.4687459 0.3606615 0.4164587
## occupied   0.6295411 0.5818624 0.5312541 0.6393385 0.5835413
```

Veja a vinheta do pacote para modelos de múltiplas ocasiões para o cálculo dos intervalos de confiança.