

Tutorial 3: Abundancia y recapturas heterogéneas*

Curso: Métodos de captura-recaptura, UNAM. Abril, 2010.

Roberto E. Munguía-Steyer

Dpto. Ecología, IB-USP, Brasil.

rmunguia.steyer@gmail.com

1. Heterogeneidad en la probabilidad de recaptura

En el tutorial pasado aprendimos a estimar la abundancia de acuerdo a las premisas del modelo M_0 . Sin embargo, existen casos en donde el supuesto de tener una probabilidad constante de captura no se cumple y existe heterogeneidad, ya sea a nivel individual, temporal y comportamental. Pensemos en una población de roedores en la cual utilizamos algún alimento como cebo para atraerlos a las trampas o cajas con las cuales los capturamos. La probabilidad de recaptura de los animales comparada con la de captura (la primera vez que caen en la trampa) podría ser mayor. Esto sucedería por el simple hecho de que caer en la trampa representa un almuerzo gratis para los roedores. La sobrestimación de las probabilidades de recaptura redundaría en una subestimación de la abundancia. Un símil sería pensar que cada individuo en la población como pelotitas dentro de una urna. Los individuos capturados y no capturados no se estarían mezclando en la urna y no serían capturados (sacados de la urna) con la misma probabilidad en la siguiente ocasión. Este modelo es el conocido como modelo comportamental M_b ¹, en el cual la probabilidad de recaptura difiere comparada con la de captura [1, 6]. Cuando la probabilidad de recaptura es menor, frecuentemente se debe al estrés ocasionado por el método de captura del animal y el tiempo que permanece en cautiverio.

La variación de las probabilidades de recaptura en el tiempo también son factibles de ocurrir. Pensemos en una población de mariposas las cuales en días lluviosos se resguardan y bajan su actividad en los sitios de forrajeo y reproducción. Lo mismo podría decirse de las variaciones de temperatura que regulan la actividad (y detectabilidad) en muchos ectotermos reptilianos (serpientes, lagartos, etc). El modelo que considera la variación temporal en las probabilidades de recaptura es conocido como M_t . La variación



* Bajo licencia: Creative Commons attribution-share alike.

¹La notación utilizada en estos modelos fue desarrollada por Otis y colaboradores y es de uso frecuente en las publicaciones académicas.

individual en las probabilidades de recaptura no atribuida a ninguna variable predictora es conocida como M_h . Cabe hacer mención que puede haber casos donde existan combinaciones en la variación de las probabilidades de recaptura mencionadas previamente. Un ejemplo de ello sería el modelo M_{bh} el cual contaría con heterogeneidad comportamental e individual [5, 6].

2. Generación de una historia de captura recaptura proveniente del modelo M_t

Usemos nuestro ejemplo de la mariposa en la cual en los días soleados $p = 0.4$ es más detectable que los lluviosos $p = 0.15$. Fuimos al campo a capturar y marcar individuos por seis días, en los cuales en el primer, tercer y cuarto día llovió. Hay 180 mariposas en la población, veamos si conseguimos detectar la variación temporal de las probabilidades de recaptura y contar con una estimación de la abundancia cercana a la real.

```
> set.seed(1234)
> k <- 6
> ni <- 180
> pt <- c(0.15, 0.4, 0.15, 0.15, 0.4, 0.4)
> celdas <- ni * k
> mat.hc <- matrix(rep(0, celdas), ni, k)
> for (i in 1:k) {
+   mat.hc[, i] <- rbinom(ni, 1, pt[i])
+ }
> mat.hc <- data.frame(mat.hc)
> names(mat.hc) <- c(paste("p", 1:k, sep = ""))
```

Mostramos las primeras historias de encuentro individuales:

```
> head(mat.hc)
  p1 p2 p3 p4 p5 p6
1  0  0  0  0  1  1
2  0  0  0  0  0  0
3  0  1  0  1  1  1
4  0  0  0  0  0  0
5  1  1  0  0  0  1
6  0  0  1  0  1  0
```

Y procedemos a remover las historias de encuentro de los individuos que nunca fueron capturados:

```
> frecuencias <- apply(mat.hc, 1, sum)
> rem0 <- which(frecuencias == 0)
> num0 <- length(rem0)
> dec <- ifelse(num0 < 1, mat.hc2 <- mat.hc, mat.hc2 <- mat.hc[-rem0,
+   ])
```

3. Modelos de abundancia: selección del modelo con mayor soporte

Usemos el paquete `Rcapture` [2] con el cual podemos estimar la abundancia de la población considerando los modelos con recapturas homogéneas M_0 así como los modelos con recapturas heterogéneas, ya sean del tipo comportamental, individual y temporal y combinaciones posibles.

```
> library(Rcapture)
> m <- closedp(mat.hc2)
> m
```

```
Number of captured units: 159
```

```
Abundance estimations and model fits:
```

	abundance	stderr	deviance	df	AIC
M0	188.8	7.7	142.17	61	249.3
Mt	184.5	7.0	58.96	56	176.1
Mh Chao	188.9	8.3	142.17	60	251.3
Mh Poisson2	187.7	9.1	142.12	60	251.3
Mh Darroch	186.4	11.7	142.11	60	251.3
Mh Gamma3.5	185.2	14.1	142.10	60	251.2
Mth Chao	185.8	7.7	58.68	55	177.8
Mth Poisson2	185.8	8.5	58.87	55	178.0
Mth Darroch	187.0	11.9	58.88	55	178.0
Mth Gamma3.5	187.7	15.5	58.90	55	178.0
Mb	268.2	54.1	128.14	60	237.3
Mbh	202.4	33.6	118.01	59	229.2

```
Note: 3 eta parameters has been set to zero in the Mh Chao model
```

```
Note: 3 eta parameters has been set to zero in the Mth Chao model
```

Podemos observar que el modelo M_t que considera variación temporal en las tasas de recaptura es el que presenta un menor valor en el Criterio de Información de Akaike (AIC) [3, 4]. El AIC se basa en el grado de ajuste de un modelo bajo métodos de máxima verosimilitud y el número de parámetros empleados en el mismo. El AIC es definido como:

$$AIC = -2 \cdot \text{Ln}(\mathcal{L}) + 2K \quad (1)$$

En el cual K es el número de parámetros y \mathcal{L} la verosimilitud del modelo. Existe entonces un compromiso entre el ajuste y la parsimonia (número de parámetros del modelo) [4]. Observamos que el modelo que se encuentra mejor soportado es el que incluye la variación temporal M_t . Sin embargo, los modelos M_{th} , que incluyen variación individual y temporal presentan valores muy cercanos al modelo M_t , por lo cual el grado de soporte de estos modelos es muy parecido. La diferencia entre los valores de AIC nos

puede indicar el grado de soporte del modelo con mejor AIC comparado con el resto de modelos, estimando el cociente de evidencia:

$$\text{Cociente de evidencia} = \frac{1}{e^{\frac{-1}{2\Delta_i}}} \quad (2)$$

Cabe hacer mención que la diferencia de los valores de AIC entre dos modelos es representada con el símbolo Δ . Si dos modelos difieren en dos unidades de AIC, el que tenga menor AIC tendría con un soporte 2.7 veces mayor [4]. Si la diferencia fuera de ocho unidades, el soporte sería 54.6 veces mayor.

Estimemos los intervalos de confianza al 95 % del modelo M_t :

```
> cpopt <- closedpCI.t(mat.hc2, m = "Mt")
> cpopt
```

```
Number of captured units: 159
```

```
Poisson estimation and model fit:
```

	abundance	stderr	deviance	df	AIC
Mt	184.5	7	58.96	56	176.1

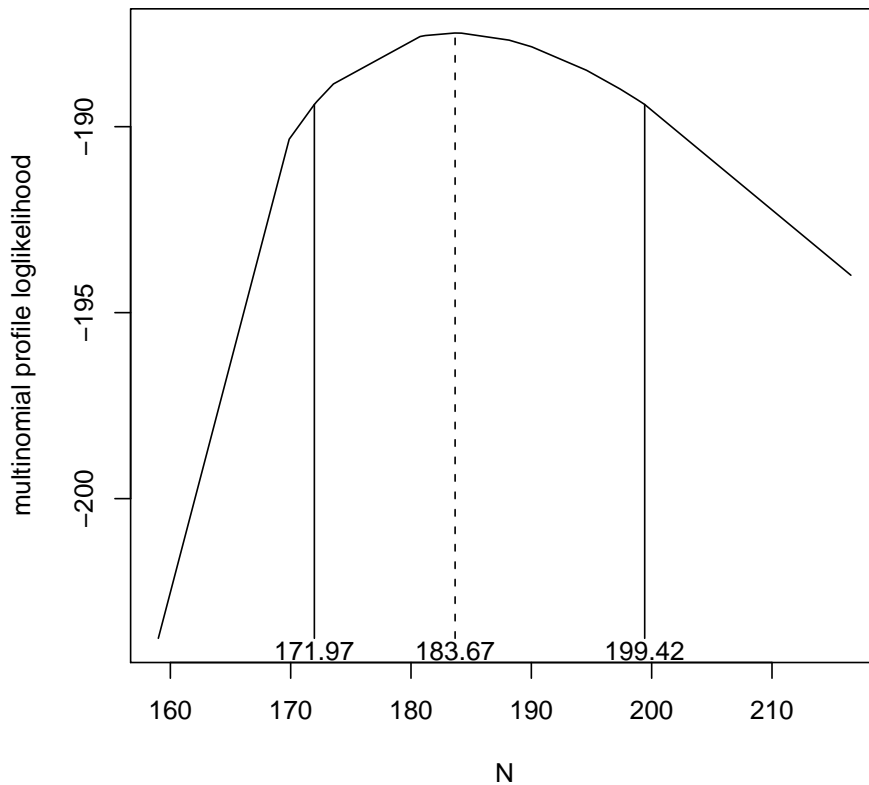
```
Multinomial estimation, 95% profile likelihood confidence interval:
```

	abundance	InfCL	SupCL
Mt	183.7	172	199.4

La abundancia estimada se encuentra próxima al valor real y el rango comprendido por los intervalos de confianza al 95 % lo incluyen. Grafiquemos el perfil de verosimilitud:

```
> plotCI(cpopt)
```

Profile Likelihood Confidence Interval



El perfil de verosimilitud muestra la típica asimetría en los intervalos de confianza debido a que la caída en los valores de verosimilitud en los valores menores a la abundancia estimada es más pronunciada. Por último, exportemos la base de datos para estimar la abundancia y la probabilidad de recaptura para diferentes modelos con capturas homogéneas y heterogéneas en el programa MARK. Usemos el formato de la historia de encuentros individuales:

```
> n.filas <- dim(mat.hc2)[1]
> hc.exp <- rep(0, n.filas)
> mat.hc2 <- as.matrix(mat.hc2)
> for (i in 1:n.filas) hc.exp[i] <- paste(mat.hc2[i, ], sep = "",
+   collapse = "")
> unos <- rep("1;", n.filas)
> expor <- cbind(hc.exp, unos)
> expor[1:6, ]

      hc.exp  unos
[1,] "000011" "1;"
[2,] "010111" "1;"
[3,] "110001" "1;"
[4,] "001010" "1;"
```

```
[5,] "010000" "1;"  
[6,] "010100" "1;"
```

```
> write.table(expor, "Mariposas.inp", quote = F, row.names = F,  
+ col.names = F, sep = " ")
```

En MARK compararemos el modelo M_t con el resto de modelos de heterogeneidad en las tasas de recaptura propuestos por Otis. Adicionalmente, construiremos un modelo que considere explícitamente la información de cuando ocurrieron los días lluviosos y soleados, especificando parámetros de recaptura diferenciados para esos días.

Referencias

- [1] S.C. Amstrup, T.L. McDonald, and B.F.J. Manly. *Handbook of capture-recapture analysis*. Princeton University Press, 2005.
- [2] S. Baillargeon and L.P. Rivest. Rcapture: Loglinear Models for Capture-Recapture in R. *Journal of Statistical Software*, 19(5):1–31, 2007.
- [3] B.M. Bolker. *Ecological models and data in R*. Princeton University Press, 2008.
- [4] K.P. Burnham and D.R. Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Verlag, 2002.
- [5] A. Chao. An overview of closed capture-recapture models. *Journal of agricultural, biological, and environmental statistics*, 6(2):158–175, 2001.
- [6] D.L. Otis, K.P. Burnham, G.C. White, and D.R. Anderson. Statistical inference from capture data on closed animal populations. *Wildlife Monographs*, 62:3–135, 1978.