**Inferencia estadística a partir de varios modelos y su utilidad en ecología**

Cayetano Gutiérrez-Cánovas1, Gema Escribano-Avila 2

Correo: [tano.gutierrez@ub.edu](mailto:tano.gutierrez@ub.edu), [gema.escribano.avila@gmail.com](mailto:gema.escribano.avila@gmail.com)

1. Freshwater Ecology, Hydrology and Management group, Deptartament de Biologia Evolutiva, Ecologia i Ciències Ambientals, Facultat de Biologia - Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Diagonal, 643, 08028 Barcelona, España.

2. Tecnigral. Consultoría Ambiental. C/. Príncipe de Vergara, 210; esc. A, 1ºD Madrid 28002 España

Cuando hacemos modelos ecológicos, normalmente nos interesa conocer la importancia y el efecto de una serie de variables potencialmente explicativas en una variable respuesta de interés. Tradicionalmente, la estrategia escogida por muchos ecólogos ha sido la de contrastar nuestro modelo con una hipótesis nula o la de añadir o quitar variables explicativas, paso a paso, hasta encontrar el “mejor” modelo (Johnson and Omland, 2004).

Sin embargo, en muchas ocasiones, existen varias hipótesis nulas o alternativas que queremos testar simultáneamente a través de varios modelos y estos métodos tradicionales no lo permiten, al tiempo que pueden ignorar variables explicativas importantes que no aparezcan en el modelo elegido (Johnson and Omland, 2004). En esta nota queremos presentar una breve introducción a la selección de modelos y a la inferencia a partir de múltiples modelos (*multi-model inference*). Estas técnicas permiten abordar las limitaciones anteriormente comentadas a través de la selección y promediado de un conjunto de modelos elegidos según su capacidad para ajustarse a los datos analizados y atendiendo al principio de parsimonia. Para profundizar en estas técnicas, recomendamos la lectura de literatura clave (Burnham and Anderson, 2002; Grueber et al., 2011), junto a las principales críticas y limitaciones (Cade, 2015; Tyre, 2017; Walker, 2018).

La aplicación de esta aproximación se diferencia del típico contraste de hipótesis, porque comparamos varios modelos alternativos a la vez, y en lugar rechazar/aceptar la hipótesis nula basándonos en la significación estadística, se valora el grado de apoyo estadístico recibido por un modelo o un grupo de modelos en base a capacidad explicativa de los datos y el principio de parsimonia. El procedimiento a seguir se muestra en la Figura 1 y será ejemplificado en esta nota usando las funciones de la librería *MuMIn* (Bartoń, 2016). Esta secuencia consta de cuatro pasos fundamentales: 1) definir el modelo global , 2) generar y ordenar los modelos en función de su capacidad explicativa y complejidad, 3) seleccionar el modelo o grupo de modelos de confianza y 4) promediar los coeficientes de las variables explicativas de estos modelos (cuando no haya un único mejor modelo).

En primer lugar, definiremos lo que se conoce como modelo global, que incluye la variable respuesta y todas las variables explicativas. De manera alternativa, también se podría construir un conjunto específico de modelos que se desee comparar sin incluir necesariamente todos los modelos. Resulta indispensable que solo incluyamos variables explicativas que sean independientes (baja colinealidad) y tengan un sentido biológico o ecológico contrastado para evitar relaciones espurias. Además, previamente, debemos estandarizar las variables explicativas para poder comparar y estimar sus tamaños de efecto de forma adecuada (Grueber et al., 2011). Para construir el modelo global usaremos la técnica que sea más apropiada para nuestros datos. En este ejemplo usaremos un modelo lineal generalizado (GLM) con error gaussiano.

**# Modelo global, función genérica del GLM gaussiano, lm()**

**mod <- lm (y ~ x1 + x2 + x3 + x4 + x5, data = dat)**

**summary(mod)**

En el segundo paso, usaremos la función *dredge()* de la librería *MuMIn* (Bartoń, 2016) para generar todos los modelos posibles a partir de distintas combinaciones de las variables explicativas contenidas en el modelo global. Seguidamente, las ordenaremos en función del índice de información de Akaike (Akaike Information Criterion, AIC), su variante para muestras pequeñas (AICc se usa cuando el número de observaciones / parámetros estimados <40; Burnham and Anderson, 2002) u otra medida análoga apropiada (Grueber et al., 2011; Johnson and Omland, 2004). El peso relativo de evidencia (*Akaike weight*) nos indica la probabilidad de que un modelo sea el mejor modelo aproximado en relación a nuestros datos, por lo que la suma del *weight* de todos los modelos siempre sumará 1. En el caso de que un modelo tenga un *weight* mayor que 0,90 bastaría para seleccionarlo como modelo final y no sería necesario realizar los pasos siguientes (selección de modelos y promedio de sus coeficientes). En esta fase debemos examinar cuidadosamente los modelos generados para evaluar su significado ecológico, en especial aquellos con un mejor balance entre la capacidad explicativa y la parsimonia del modelo (AIC más bajo).

**# ejecuta todos los modelos posibles y clasifica los modelos en función del AICc de cada modelo. También se visualiza la bondad de ajuste**

**options(na.action = "na.fail") # necesario para ejecutar dredge() puesto que no se pueden comparar modelos con distinto número de casos**

**mod\_d <- dredge (mod, rank = "AIC", extra = c(“R^2”)))**

**mod\_d # ranking de los modelos producidos**

En tercer lugar, cuando no hayamos identificado un único mejor modelo, usaremos la función *get.models()* para seleccionar el grupo de modelos de confianza en función de su AIC o de su *weight*. Con esta función, podemos seleccionar aquellos modelos con una diferencia en el AIC (delta o ΔAIC) igual o inferior a 2, 4 o 6 unidades respecto al modelo con el AIC menor, o seleccionar el mínimo número de mejores modelos cuyo *weight* conjunto sea ≥ 0,95. Éste será nuestro grupo de modelos de confianza. A día de hoy, no existe consenso sobre qué método seguir (Grueber et al., 2011), así que dependerá de la pregunta ecológica a responder. También podemos calcular la bondad de ajuste (*r*2) de cada modelo usando, por ejemplo, la función *r.squaredGLMM()* (Nakagawa and Schielzeth, 2013), que es recomendable si estamos usando modelos mixtos. De esta manera podremos saber si el grupo de modelos de confianza presentan un buen ajuste a los datos, lo cual es importante a la hora de interpretar los resultados obtenidos como comentaremos más adelante. Además, debemos evaluar las asunciones de cada modelo (e.g. normalidad y homocedasticidad de residuos) y el rango de variación y la distribución de los tamaños de efecto (coeficientes de regresión) para cada variable explicativa en los modelos de confianza para asegurarnos de que el valor medio del coeficiente es representativo.

**# ejemplo1: subconjunto delta AIC ≤ 2**

**mod\_set <- get.models(mod\_d, subset = delta <= 2)**

**# ejemplo2: subconjunto AIC ≤ 0.95**

**mod\_set <- get.models(mod\_d, subset = cumsum (mod\_d$weight) <= 0.95)**

Finalmente, podremos utilizar la función *model.avg()* para promediar los coeficientes del grupo de modelos de confianza, cuando no haya un único mejor modelo. Así, obtendremos los coeficientes (tamaño del efecto) promedio y el valor *p* para cada una de las variables explicativas contenidas en los modelos seleccionados. Esta función nos ofrece dos tipos de promediado que, en ambos casos, proporcionan una media de los coeficientes ponderada por el *weight* de cada modelo. El *zero-method average* (*full average* en R) considera el valor de los coeficientes de cada variable explicativa que aparecen en los modelos seleccionados y el valor cero para aquellos modelos que no contengan dicha variable. Por el contrario, el *natural average* (*conditional average* en R) solo promedia los valores de los coeficientes incluidos en los modelos seleccionados. Aunque no existe un consenso sobre cuál de los dos métodos es más apropiado (Grueber et al., 2011), el primero es más conservador. Por lo tanto, la elección de uno u otro dependerá de la pregunta a resolver. En cualquier caso, es importante tener presente que un grupo de modelos de confianza con *r*2 muy bajos arrojará necesariamente coeficientes poco precisos y, por lo tanto, tanto los modelos de confianza como el modelo promediado tendrán una mala capacidad predictiva.

**mod\_av <- model.avg(mod\_set, revised.var = TRUE) # model averaging**

**summary(mod\_av) # resultados del multi-model averaging**

De manera adicional, también podemos comparar la media de las predicciones de los modelos de confianza y con las del modelo con coeficientes promediados (*zero-method* y *natural average*). Aunque la inferencia a partir de varios modelos nos puede ayudar a reducir la incertidumbre en el modelado ecológico, existe un debate activo sobre sus limitaciones reales y las circunstancias en las cuales podemos aplicarla de forma segura (Banner and Higgs, 2017; Tyre, 2017). En cualquier caso, debemos de ser cuidadosos en la selección de variables explicativas y evitar la tentación de usar esta técnica como un sistema automático (no supervisado) de selección de modelos.

**Agradecimientos**

Gracias al grupo de Ecoinformática de la AEET por la revisión de esta nota

**Referencias**

Banner, K.M., Higgs, M.D., 2017. Considerations for assessing model averaging of regression coefficients: Ecol. Appl. 27, 78–93. doi:10.1002/eap.1419

Bartoń, K., 2016. MuMIn: Multi-model inference. R package version 1.15.6. Version 1, 18.

Burnham, K.P., Anderson, D.R., 2002. Model selection and multimodel inference: A practical information-theoretic approach (2nd ed), Ecological Modelling. doi:10.1016/j.ecolmodel.2003.11.004

Cade, B., 2015. Model averaging and muddled multimodel inferences. Ecology 96, 2370–2382.

Grueber, C.E., Nakagawa, S., Laws, R.J., Jamieson, I.G., 2011. Multimodel inference in ecology and evolution: Challenges and solutions. J. Evol. Biol. 24, 699–711. doi:10.1111/j.1420-9101.2010.02210.x

Johnson, J.B., Omland, K.S., 2004. Model selection in ecology and evolution. Trends Ecol. Evol. 19, 101–108. doi:10.1016/j.tree.2003.10.013

Nakagawa, S., Schielzeth, H., 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol. Evol. 4, 133–142. doi:10.1111/j.2041-210x.2012.00261.x

Tyre, D., 2017. Does model averaging make sense? [WWW Document]. A few cheap shots. URL https://atyre2.github.io/2016/09/03/sum-to-zero-contrasts.html

Walker, J.A., 2018. On model averaging partial regression coefficients. bioRxiv. doi:10.1111/j.1468-5922.2007.00690.x