**Ventajas de la estadística bayesiana frente a la frecuentista: ¿por qué nos resistimos a usarla?**

Paloma Ruiz-Benito1, Enrique Andivia1, Juliette Archambeaou2, Julen Astigarraga1, Rafael Barrientos1, Verónica Cruz-Alonso1, Daniel Gómez1, Loreto Martínez-Baroja1, Pablo Quiles1, Zoë Roher1, Ana M.C. Santos1, Elena Velado1, Sara Villén-Pérez1, Ignacio Morales-Castilla1

(1) Grupo de Ecología y Restauración Forestal, Departamento de Ciencias de la Vida, Universidad de Alcalá, Edificio de Ciencias, Campus Universitario, 28805 Alcalá de Henares (Madrid), España.

(2) UMR 1202 BIOGECO, University of Bordeaux, Talence, France

Autor de correspondencia: P. Ruiz-Benito [palomaruizbenito@gmail.com]

La complejidad de la naturaleza requiere de la aplicación de modelos matemáticos que nos permitan su simplificación, es decir, que nos permitan realizar abstracciones para describir nuestras observaciones (Bolker et al. 2008). En última instancia, todos los modelos son erróneos, pero algunos son útiles para responder a determinadas preguntas. La teoría de la probabilidad permite la descripción de la incertidumbre asociada a los sistemas estocásticos naturales y el ajuste de modelos. Para describir la realidad y testar nuestra hipótesis calculamos una serie de parámetros asociados al proceso objeto de estudio, que generalmente incluyen al menos un estadístico de tendencia central que lleva asociado un rango de variación (p. ej. media e intervalos de confianza si usamos una aproximación frecuentista, o la media de la distribución posterior de parámetro y los intervalos de credibilidad en estadística bayesiana). Los tres enfoques principales usados para el cálculo de estos parámetros son la estadística frecuentista, la máxima verosimilitud y la estadística bayesiana (Tabla 1). Estos enfoques se diferencian principalmente en los métodos utilizados para extraer conclusiones sobre los datos observados: los p-valores, los estimadores de máxima verosimilitud y la distribución posterior de los parámetros.

Los p-valores y los estimadores de máxima verosimilitud, comúnmente utilizados, han sido criticados porque algunas asunciones no se ajustan al funcionamiento de los procesos ecológicos y precisan de normas arbitrarias e inflexibles para aceptar o rechazar hipótesis (Wasserstein y Lazar, 2016). Por ejemplo, en muchos estudios se acepta la hipótesis de que dos grupos de datos son diferentes cuando el p-valor asociado al estadístico de contraste es < 0.05, es decir, cuando hay una probabilidad máxima del 5% de que las diferencias entre las distribuciones de esos dos grupos de datos y, por tanto, la obtención de ese estadístico sea consecuencia del azar. Además, el p-valor está muy afectado por el tamaño muestral y su interpretación testando la hipótesis nula puede resultar poco intuitiva (Anderson et al. 2000). Frente al p-valor, los criterios de información como el AIC (Akaike Information Criterion), BIC (Bayesian Informatión Criterion) o DIC (Deviance Information Criterion) permiten realizar la comparación y selección de modelos testando simultáneamente varias hipótesis específicas (Johnson y Omald 2004). Los Criterios de Información se basan en el principio de parsimonia de forma que se evalúa la verosimilitud del modelo en función de su bondad de ajuste a los datos y de su complejidad (número de parámetros, Burnham y Anderson, 2002). La selección de modelos se hace generalmente en función de diferencias en AIC entre modelos (cuando ΔAIC > 2 se considera que hay soporte empírico para el modelo más complejo, Burnham y Anderson, 2002), y la inferencia de múltiples modelos permite calcular la media de todos ellos cuando no hay un grado de soporte empírico claro para un determinado modelo (Dorman et al. 2018). A pesar de la popularidad creciente de este tipo de enfoques, también han sido criticados dado que, si bien serían útiles para descartar los modelos menos informativos, no bastarían para identificar el mejor modelo (Kadane y Lazar 2004), para lo que serían necesarios criterios adicionales (Dziak et al. 2017).

La estadística bayesiana es el único enfoque en el que se hace un uso explícito de la probabilidad para cuantificar la incertidumbre de la inferencia (Tabla 1). Se trata de un proceso de aprendizaje iterativo en el que se alcanzan conclusiones sobre un fenómeno (probabilidad a posteriori) a partir del conocimiento previo sobre el sistema (probabilidad a priori) y de nuevas evidencias (información proveniente de los datos). Es decir, los resultados de un nuevo estudio podrían ser usados para actualizar el conocimiento sobre el sistema e incluirlo en estudios posteriores. Por tanto, la estadística bayesiana se basa en la reasignación de credibilidad a través de las posibilidades, donde las posibilidades son los valores de los parámetros en modelos matemáticos biológicamente coherentes (Kruschke 2015). Mientras que la estadística frecuentista se basa en una lógica deductiva, la estadística bayesiana trabaja con una lógica inductiva en la que las hipótesis no son testadas, sino que se comprueba su credibilidad en base a la evidencia empírica.

El uso de un marco de trabajo bayesiano tiene una serie de ventajas sobre la estadística frecuentista (p. ej. Bolker et al. 2004; Gelman et al. 2004): (i) permite ajustar modelos complejos no abordables por métodos frecuentistas debido a las restricciones de estos modelos (p. ej. normalidad en la distribución de los parámetros); (ii) permite alcanzar estimaciones más exactas de los parámetros cuando el tamaño muestral es pequeño; (iii) la interpretación de los resultados es fácil y directa ya que indican la probabilidad de que un parámetro tome un cierto valor; (iv) se pueden incluir medidas de incertidumbre, datos perdidos y diferentes niveles de variabilidad; (v) permite realizar propagaciones de error; (v) permite especificar las distribuciones de los parámetros (dependientes a su vez de otros parámetros) cuando *a priori* sabemos cómo se distribuyen (*priors*); (vii) minimiza el uso de límites arbitrarios para tomar decisiones.

A pesar de las múltiples ventajas, la estadística bayesiana tan sólo recientemente comienza a implementarse con cierta frecuencia en el área de Ecología. Hemos realizado una búsqueda sencilla en Web of Science incluyendo las palabras “Bayesian” & “Ecology” y hemos obtenido 14495 resultados de los que el 90,7% se han publicado a partir del año 2000 y el 73,6% desde el año 2010 (Figura 1a). Las principales áreas de publicación son: Ciencias Ambientales, Zoología, Genética y Ciencias de la Vida (Figura 1b). La utilización incipiente de la estadística bayesiana en Ecología se debe a diversos motivos. En primer lugar, a la inercia en el uso de métodos frecuentistas y de máxima verosimilitud. De hecho, los resultados obtenidos por ajustes bayesianos son similares a los obtenidos mediante estadística frecuentista si no se usa un conocimiento *a priori* de los parámetros o modelos complejos. En segundo lugar, los procedimientos bayesianos son computacionalmente complejos y pueden requerir un elevado tiempo de ejecución para ajustar el modelo, especialmente si el tamaño muestral es elevado. Esta limitación se está viendo subsanada gracias al aumento de las capacidades de computación en paralelo de los equipamientos informáticos y a la mejora de paquetes de software crecientemente asequibles para usuarios no experimentados. A este respecto, cabe destacar el reciente desarrollo de programas informáticos de código abierto tales como Stan (www.mc-stan.org), que hacen uso de algoritmos de muestreo más eficientes (p.ej. Hamiltonian Monte Carlo propuesto por Hoffman y Gelman, 2014) que los utilizados por anteriores aplicaciones para la estadística bayesiana (p.ej. algoritmo de muestreo de Gibbs propuesto por Geman y Geman (1984) y utilizado por programas como Wingbugs, www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-winbugs/, y JAGS2, http://mcmc-jags.sourceforge.net/). El incremento de la eficiencia de los algoritmos, la implementación en R mediante paquetes específicos para el uso de Stan como “rstan”, o “brms”, o la posibilidad de editar scripts de Stan directamente desde RStudio, están propiciando un rápido aumento de la popularidad de la estadística bayesiana, también en el ámbito de la Ecología.

Debido a la dificultad de enfrentarnos en solitario a los retos que supone aplicar marcos bayesianos al estudio de los procesos ecológicos, hemos generado un grupo de trabajo sobre estadística bayesiana. Desde el pasado mes de marzo nos reunimos quincenalmente para realizar juntos ejemplos, solucionar problemas que encontramos al aplicar los modelos bayesianos a nuestros datos y compartir puntos críticos a conocer a la hora de realizar un modelo.

**Agradecimientos**

Gracias al grupo de Ecoinformática de la AEET por el apoyo para escribir esta nota. P.R-B. cuenta con una ayuda Atracción de Talento de la Comunidad de Madrid (2016-T2/AMB-1665). V.C-A. está financiada por el programa de Becas Predoctorales de Medio Ambiente de la Fundación Tatiana Pérez de Guzmán el Bueno (2015). A.M.C.S. cuenta con una beca Juan de la Cierva (MINECO, IJCI-2014-19502). S.V-P cuenta con una ayuda Atracción de Talento de la Comunidad de Madrid (2017-T2/AMB-6035). I.M-C. está financiado por una Ayuda Postdoctoral del Programa Propio de la Universidad de Alcalá.

**Referencias**

Anderson, D.R., Burnham, K.P., Thompson, W.L. 2000. Null hypothesis testing: problems, prevalence, and an alternative. *The Journal of Wildlife Management* 64: 912-923.

Bolker, B. 2008. Ecological models and data in R. Princeton University Press. Princeton, New Jersey, USA.

Burnham, K.P., Anderson, D.R., 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer-Verlag, New York.

Dziak, J.J., Coffman, D.L., Lanza, S.T., Li, R. 2017. Sensitivity and specificity of information criteria. *PeerJ PrePrints*: e1103v1103.

Dormann, C.F., Calabrese, J.M., Guillera-Arroita, G., Matechou, E., Bahn, V., Barto´n, K., Beale, C.M., Ciuti, S., Elith, J., Gerstner, K., Guelat, J., Keil, P., Lahoz-Monfort, J.J., Pollock, L.J. Reineking, B., Roberts, D.R., Schröder B., Tuiller, W., Warton, D.I., Wintle, B.A., Wood, S.N., W¨uest, R.O, Hartig, F. 2018. Model averaging in ecology: a review of Bayesian, information-theoretic and tactical approaches for predictive inference. *Ecological Monographs*.

Edwards, M.B. 1992. Likelihood. Johns Hopkins University Press. Baltimore, Maryland, USA.

Geman, S., Geman, D. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6: 721–741.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. Bayesian data analysis. Chapman & Hall/CRC Press, New York.

Hoffman, M.D., Gelman, A. 2014. The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research* 15: 1593-1623.

Kadane, J.B., Lazar, N.A. 2004. Methods and criteria for model selection. *Journal of the American Statistical Association* 99: 279-290.

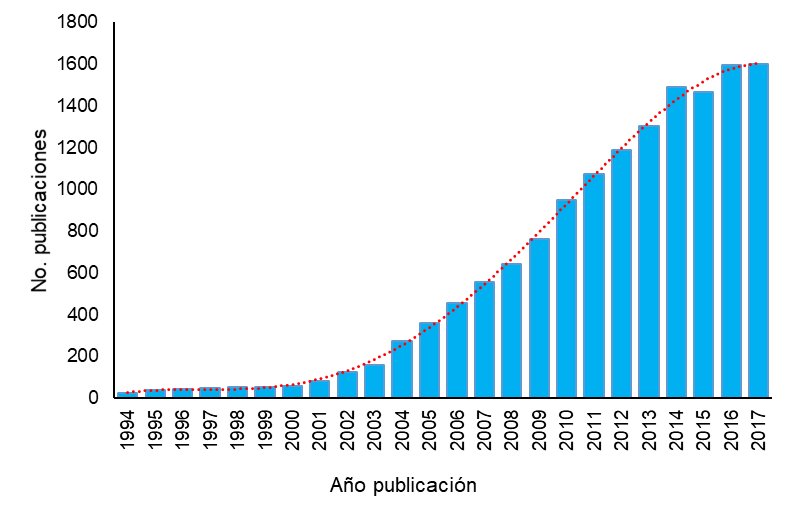
Kéry, M. 2010. Introduction to WinBUGS for Ecologists. Academic Press. Burlington, Massachusetts, USA.

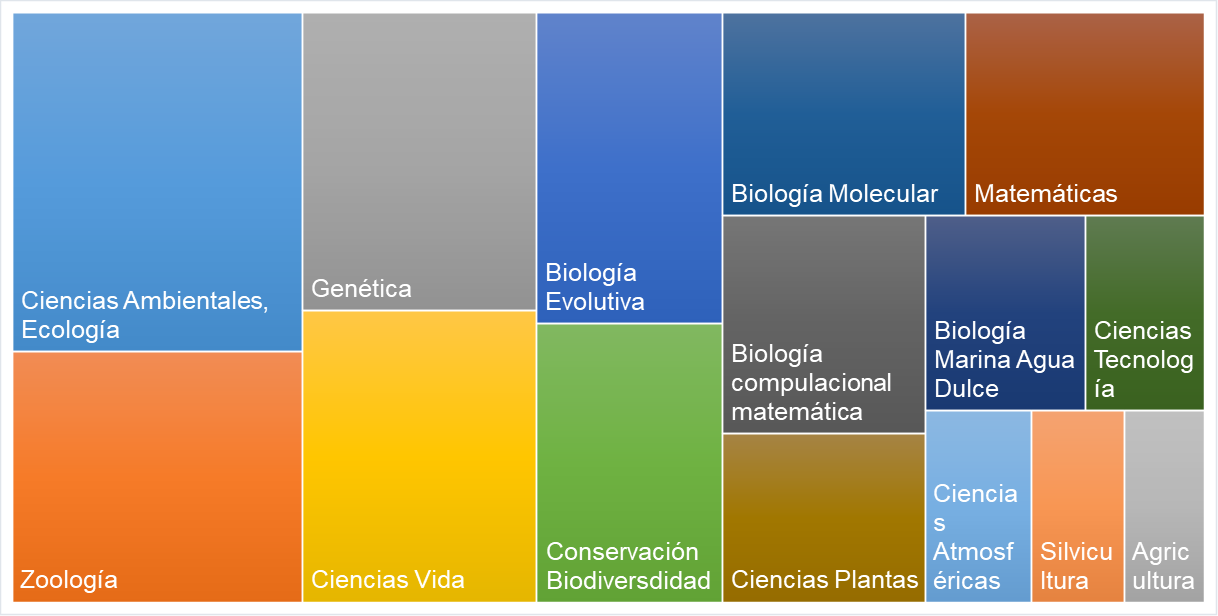
Kruschke, J.K., 2015. Doing Bayesian data analysis: a tutorial with R, JAGS, and Stan. Elsevier, USA.

Wasserstein, R. L., Lazar, N.A. 2016. The ASA's statement on p-values: context, process, and purpose. *The American Statistician*: 70, 129-133.

**Tabla 1.** Comparación de las aproximaciones frecuentistas, de máxima verosimilitud y bayesiana para ajustar un determinado modelo (p.ej. Y = a + b X).

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Frecuentista** | **Máxima verosimilitud** | **Bayesiana** |
| Aproximación probabilidad | Contraste de hipótesis basado en el p-valor: probabilidad (*P*) de observar los datos (*D*) dado que la hipótesis nula (*Ho*) es verdadera. | Probabilidad de observar los datos dado un determinado modelo (con parámetros θ) viene determinado por la verosimilitud (*L: likelihood*). | Probabilidad de que la hipótesis sea verdadera dados los datos observados. Se basa en el teorema de Bayes: |
| Parámetro | P-valor es la probabilidad de observar los datos bajo la hipótesis nula (¡no es la probabilidad de que la hipótesis nula sea verdadera!) | Parámetros que hacen que el resultado sea más probable (MLE: estimadores de máxima verosimilitud). No hay definición de probabilidad de los parámetros, pero se definen los intervalos de confianza con dos unidades de probabilidad logarítmica (Edward 1992) | Media de la distribución posterior de los parámetros. El parámetro tiene una distribución de probabilidad: esta es la única aproximación con la que podemos estimar la probabilidad de diferentes hipótesis o valores del parámetro. |

****

****

**Figura 1.** a) Número de publicaciones en la base de datos Web of Science con las palabras clave “Bayesian” & “Ecology” a lo largo del tiempo. b) Principales áreas temáticas en que se dividen las publicaciones de la búsqueda anterior.