

Lo que *es*, lo que *podría ser* y el análisis e interpretación de los datos de un estudio de campo

What *is*, what *might be*, and the analysis and interpretation of field data

Peter Feinsinger

Wildlife Conservation Society
Pasaje El Astillero N° 8, (4401) Vaqueros, Salta, Argentina
E-mail: peter.feinsinger@nau.edu

El ecólogo que pretende seguir una u otra versión del método hipotético deductivo propone una hipótesis de trabajo general y de ella deduce una predicción particular que, planteada debidamente, precisa “Si la hipótesis es válida y si mi estudio delimitado en el espacio y el tiempo cumple con sus precondiciones, entonces encontraré una relación tal entre Y y X .” Asimismo, el ecólogo que pretende seguir el ciclo de la indagación (Fig. 1; Feinsinger 2004, Feinsinger *et al.* 2010) plantea un concepto de fondo general prendido por su observación de un fenómeno en su entorno, una inquietud particular (a menudo el concepto de fondo a escala del mismo entorno) y una bien planteada pregunta de trabajo que precisa “En el espacio tal y el periodo de tiempo tal, ¿cómo varía Y entre las unidades de respuesta i de las diferentes clases (niveles) X ?” Entonces cada ecólogo diseña el estudio que mejor pruebe la predicción o mejor conteste la pregunta, respectivamente. Una vez precisado y afinado el diseño, toma los datos; luego los analiza, descubriendo las tendencias generales entre ellos y las excepciones llamativas. Ahora reflexiona sobre el significado biológico, las posibles causas y extrapolaciones e implicaciones de los hallazgos. Finalmente el ecólogo redacta un manuscrito para la tesis o para la publicación científica, consistiendo en la introducción (= planteamiento de la pregunta del ciclo de la indagación, Fig. 1), métodos, resultados (= la acción) y la discusión (= la reflexión).

Lo que *es* y su análisis estadístico

Al terminar el trabajo de campo el ecólogo vuelve a su institución con miles de datos tomados fielmente según el diseño del estudio. El conjunto de N datos constituye *lo que es* (técnicamente, lo que era) según aquel diseño y el azar del muestreo: Cada unidad de respuesta i está caracterizada por un valor o registro de Y , sea que haya submuestreado (tomado más de un dato por unidad de respuesta a fin de darle un valor más confiable de lo que está midiendo a la unidad como un todo) o no. Los datos obtenidos constituyen la respuesta directa a la pregunta original y deben ser presentados en un gráfico o tabla. Sin embargo, con el gráfico o tabla al ecólogo le podría ser difícil ver una tendencia o la falta de ésta. Entonces, emprende el *análisis estadístico* para resumir, entender mejor y presentar más claramente los resultados: Calcula *estadísticos de la(s) muestra(s)*, como la media aritmética más la desviación estándar. Sin embargo, debería reconocer que al resumir las tendencias, ya no está contestando directamente su pregunta inicial (que incluía todos los datos tomados, no los promedios ni grados de variación en promedio).

Hay un sinnúmero de análisis estadísticos, desde los más sencillos (p.e. la media aritmética y la desviación estándar) hasta los más complejos (como la estadística multivariante). Pero ¡cuidado! los análisis estadísticos, hasta los usados con mayor frecuencia, pueden resumir excesivamente,

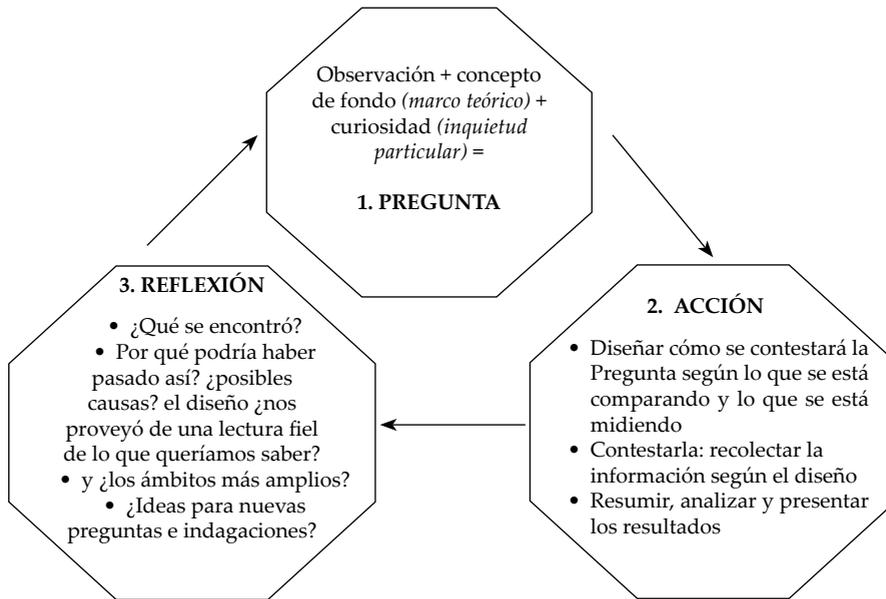


Figura 1. El ciclo de la indagación, modificado ligeramente de Feinsinger *et al.* (2010).

distorsionar y hasta tergiversar *lo que es*. El simple cálculo de la media aritmética y la varianza puede resumir excesivamente. El simple análisis de regresión lineal (y más todavía las regresiones múltiples y análisis GLM) puede distorsionar: ¿cuántos fenómenos ecológicos en el campo presentarían una relación lineal entre Y y X ? ¿No sería mejor aplicar análisis estadísticos que no asuman una relación lineal entre lo que se compara y lo que se mide (ver Keele 2007)? El simple índice de diversidad de especies - una forma particular de estadístico de la muestra - tergiversa por completo los datos de la biodiversidad (Feinsinger 2004: cap. 9). Lo menos que se puede decir de los índices de diversidad de especies es que pierden la información de mayor significado biológico: Las identidades de las especies y su abundancia relativa. Así, los análisis estadísticos de *lo que es* (los resultados) pueden ser útiles y a menudo necesarios, pero también pueden ser peligrosos. ¿La solución? Presentar bien los datos originales según el diseño, un dato por cada unidad de

respuesta; y no dejar de emplear el análisis estadístico, pero cuestionar la relevancia biológica de cada alternativa.

Lo que *podría ser* I: La inferencia estadística

¿Nos quedamos con los datos tomados? o ¿deberíamos reconocer la posibilidad en que la tendencia o falta de tendencias entre ellos sea ilusoria? Es imprescindible que reconozcamos esa posibilidad y teóricamente pueda ser evaluada cuantitativamente mediante la *inferencia estadística*. Esta es la manera de aprovechar los datos tomados (*lo que es*) y los estadísticos de la muestra ya calculados para inferir hacia *lo que podría ser*: los datos tomados junto a los no tomados, los que en teoría podrían obtenerse si pudiéramos seguir muestreando bajo las mismas condiciones que produjeron los datos tomados.

Es muy probable que la clase de inferencia estadística que un ecólogo emplea sea una

prueba. Las pruebas estadísticas se refieren al “NHST”, el probar la significación de la hipótesis nula (en inglés: null hypothesis statistical testing). La inquietud al fondo del NHST es ¿cuál es la probabilidad de que las tendencias mostradas por los datos tomados o *tendencias más llamativas todavía*, podrían obtenerse por muestrear al azar una población estadística de datos sin ninguna tendencia entre ellos? Esa probabilidad se llama P . Tradicionalmente dejamos que el valor de P decida el significado (la significación) de la tendencia entre los datos tomados: $P \leq 0.05 =$ significativa, $P > 0.05 =$ no significativa.

Sin embargo, el NHST ya está desacreditado entre la mayoría de los estadísticos y científicos no biólogos (Cohen 1994, Hubbard & Lindsay 2008, Silva-Ayçaguer *et al.* 2010) y entre cada vez más ecólogos y biólogos de la conservación (Johnson 1999, Fidler *et al.* 2005, Estay & Naulin 2011). El NHST resulta principalmente de una sola persona, R. A. Fisher (Hurlbert & Lombardi 2009). Hay numerosas críticas fuertes del mal uso, abuso y uso excesivo del NHST (Fidler *et al.* 2006, Lombardi & Hurlbert 2009). Y ¿qué quiere decir un valor de P , ya que (a) refleja más que nada el tamaño de la muestra, (b) no habla de la magnitud del efecto de X sobre Y y (c) por lo tanto no tiene relación con el significado biológico (Feinsinger 2004)?

El simple cálculo de los *intervalos de confianza* basado en los estadísticos de la muestra y la comparación de los intervalos de confianza entre muestras es mucho más fácil de interpretar y aplicar que las pruebas estadísticas (Nakagawa & Cuthill 2007). Pero los intervalos de confianza en sí no son una panacea (Hurlbert & Lombardi 2009). Sus cálculos, como los de las pruebas paramétricas del NHST, incluyen supuestos matemáticos con los que nuestros datos ecológicos casi nunca cumplen. El significado biológico de las comparaciones de los intervalos de confianza sigue eludiéndonos, aunque no tanto como el del NHST. Sin embargo, el cálculo de intervalos de confianza sobre la *magnitud del efecto* (Cohen

1988, Nakagawa & Cuthill 2007) sí habla directamente del significado biológico de los datos tomados más los no tomados). De todos los acercamientos de la inferencia estadística tradicional o “frecuentista”, la estimación de intervalos de confianza sobre la magnitud del efecto según las técnicas de remuestreo es la más confiable y menos desacreditada (Mielke & Berry 2001, D. Denis Ávila 2012, com. pers.).

Enteoría los acercamientos “nofrecuentistas” a la inferencia estadística también hablan directamente del significado biológico. La estadística bayesiana, el criterio de información de Akaike y otros modelos teóricos de información (en inglés: IT models, Stephens *et al.* 2007) están cada vez más usados en la ecología y la conservación biológica (Fidler *et al.* 2006). Sin embargo, despiertan nuevas dudas (Hurlbert & Lombardi 2009). Sin que el investigador conozca bien la historia natural de lo que estudia, su selección *a priori* de modelos (paso necesario para la mayoría de estos acercamientos) puede carecer de relevancia. Los acercamientos incluyen supuestos matemáticos con los cuales nuestros datos casi nunca cumplen y su precisión y exactitud dependen notoriamente del tamaño de la muestra; pero estos hechos son menos reconocidos, evidentes y más fáciles de ignorar que en la inferencia frecuentista. La interpretación biológica de los resultados no es fácil. La observación que los acercamientos no frecuentistas están de moda no significa que sean confiables a la hora de aplicarlos a los datos de campo.

Debemos reconocer que las tendencias entre los datos tomados podrían representar bien o mal las tendencias reales, pero desconocidos, a través de los datos tomados sumados a los no tomados. El propósito de la inferencia estadística es admirable: Provee las herramientas para aprovechar los datos tomados y poner probabilidades sobre *lo que podría ser*. Pero los inventores de las herramientas – quienes discuten mucho entre sí – suponen que los usuarios de sus inventos hayan tomado muchos datos al azar con

respecto a una población estadística de datos que refleje la pregunta original y satisfaga los requisitos matemáticos; además que conozcan bien la manera de emplear debidamente las herramientas. Muy pocos trabajos en ecología cumplen con esos supuestos. Es como si los ecólogos fuéramos niños preescolares y los inventores de las herramientas, las maestras. Sin malas intenciones, ellas abrieron la puerta de un salón llena de motosierras prendidas, nos dijeron “¡Jueguen niños!” y salieron. ¿El resultado? Adivinen. ¿La solución? Diseñar el estudio más fuerte y riguroso que sea posible, para que *lo que es* (los datos tomados) nos dé una lectura fiel *con poca duda* de cómo varía lo que medimos (Y) entre lo que comparamos (las unidades de respuesta *i* de las diferentes clases X) y *quedarnos con eso, sin abrir la puerta del salón de las motosierras* - o formarnos bien como forestales cuidadosos a fin de poder aplicar debidamente las motosierras a *lo que podría ser*.

Lo que *podría ser* II: La Reflexión

La reflexión (Fig. 1) trata sobre otras perspectivas de *lo que es* y *lo que podría ser*. Al comienzo de la reflexión (o primer párrafo de la discusión del manuscrito), miramos los resultados desde cierta distancia y nos preguntamos ¿cuáles son las tendencias entre los datos tomados y las excepciones llamativas de esas tendencias? Sobre esa base especulamos. Primero nos preguntamos ¿por qué podría haber resultado así? Debemos reconocer que *lo que es* ya terminó, al final de aquella mirada de los resultados. Ahora estamos hablando de *lo que podría ser* o más precisamente de *lo que podría haber sido* (causas explicativas de los resultados). Debemos cambiar los verbos indicativos o ciertos por los verbos condicionales o con algún grado de incertidumbre (“podría ser que”, “podría haber sido que”, “podría deberse a”, “proponemos que”, “es posible que”, “es probable que”, “es casi cierto que”, pero **nunca** “es” o “se debe a”). Sabemos sólo los datos tomados dentro del

ámbito espacial y temporal del muestreo. No hemos comprobado las causas responsables y no podemos extrapolar a espacios o tiempos fuera de aquel ámbito como si fuera “es así”.

En fin, como dijo el tutor de Simón Bolívar “*hay razón de dudar de toda aserción que no sea el resultado de un trabajo consumado*” (Rodríguez 1840). El problema es nuestra tendencia humana de aseverar, afirmar y generalizar excesivamente. La solución: Diferenciar claramente entre *lo que podría ser* y *lo que es* por el uso adecuado de los verbos; y aprovechar *lo que podría ser* para cumplir con la última fase de la reflexión (Fig. 1), especulando creativamente sobre lo que podría pasar bajo otras condiciones y proponiendo nuevas indagaciones.

Lo que *podría ser* III: La modelación

Algunos resultados y algunas especulaciones de la reflexión pueden someterse a la modelación. En los trabajos de la ecología, el manejo de fauna y muchos otros campos se encuentra cada vez más modelación. Unas modelaciones aprovechan los datos tomados (*lo que es*) para tratar aquella última frase de la reflexión: ¿qué podría suceder si cambiáramos las condiciones? Ejemplos incluyen la modelación de los efectos de cambios climáticos sobre la distribución y supervivencia de especies particulares o sobre la composición de la biota como un todo. Otras modelaciones aprovechan los datos tomados (*lo que es*) para predecir lo que podría ser si pudiéramos hacer censos completos, por ejemplo la modelación de la rarefacción para estimar la riqueza de especies S o la de la ocupación de hábitats (en inglés: habitat occupancy models). Todos los modelos necesariamente incluyen simplificaciones y no sólo supuestos matemáticos (como es la inferencia estadística), sino también biológicos. Los resultados de la modelación no son resultados del estudio, sino son especulaciones cuantitativas sobre lo que podría suceder si nuestro muestreo inicial o las condiciones bajo las cuales lo realizamos fueran diferentes. Así

casi siempre pertenecen a la fase de reflexión (Fig. 1) o la discusión del manuscrito.

La modelación está de moda. Es más barato, rápido, cómodo y fácil modelar que pasar años en el campo tomando abundantes datos según un diseño riguroso y descubriendo las sorpresas y complejidades de la historia natural (Lindenmayer & Likens 2011). Es imprescindible para el manejo adaptativo (en inglés: adaptive management) y otros campos, pero debemos reconocer que la modelación es sólo la reflexión cuantitativa y que la calidad y confiabilidad de los resultados (*lo que es*) depende de las de los datos originales ingresados en los programas de computación. Los inventores y primeros usuarios de la modelación tomaron bien numerosos datos (ver el prefacio de MacKenzie *et al.* 2006). Pero por ignorancia o porque quieren publicar rápido, muchos usuarios posteriores toman pocos datos y/o toman mal los datos. El programa de modelación funciona igual y conduce a resultados cuantitativos e impresionantes, pero ¿qué significan? La modelación es más joven que la inferencia estadística, pero es más susceptible todavía al uso excesivo, mal uso y abuso. Las consecuencias del uso indebido pueden ser graves, ya que se emplea mucho en el manejo y la conservación. Tengamos cuidado.

En resumen, no confundamos *lo que es* con *lo que podría ser* en lo que calculamos, escribimos y pensamos, menos todavía en lo que aplicamos al manejo y la conservación. Recordemos las palabras de un cubano naturalista (entre otras cosas) del siglo XIX, José Martí (Váldez Galárraga 2004): “Es un presumido el que se crea más sabio que la naturaleza” y “¡Qué tristeza, ver tanto y saber tan poco!”.

Agradecimientos

Agradezco a M. Moraes, M. Nuñez-Regueiro y I. Ventosa los comentarios sobre el manuscrito y a D. Denis, A. Tálamo y C. Trucco por las discusiones que han contribuido mucho a las perspectivas presentadas aquí. Dedico el

editorial a R. B. Root, mentor sin par, gran naturalista y gran ecólogo que está para despedirse de nosotros y pasar de “lo que es” a “lo que podría ser”.

Referencias

- Cohen, J. 1988. Statistical power analysis for the behavioral sciences, 2^{da} ed. Lawrence Erlbaum Associates, Hillsdale, Nueva Jersey. 567 p.
- Cohen, J. 1996. The earth is round ($p < .05$). *American Psychologist* 49: 997-1002.
- Estay, S. A. & P. I. Naulin. 2011. Data analysis in forest sciences: why do we continue using null hypothesis significance tests? *Bosque* 32: 3-9.
- Feinsinger, P. 2004. El diseño de estudios de campo para la conservación de la biodiversidad. Editorial FAN, Santa Cruz. 242 p.
- Feinsinger, P., C. Pozzi, C. Trucco, R. L. Cuéllar, A. Laina, M. Cañizares & A. Noss. 2010. Investigación, conservación y los espacios protegidos de América latina: una historia incompleta. *Ecosistemas* 19(2). <http://www.revistaecosistemas.net/articulo.asp?Id=645>
- Fidler, F., M. A. Burgman, G. Cumming, R. Buttrose & Neil Thomason. 2006. Impact of criticism of null-hypothesis significance testing on statistical reporting practices in conservation biology. *Conservation Biology* 20: 1539-1544.
- Hubbard, R. & R. M. Lindsay. 2008. Why *P* values are not a useful measure of evidence in statistical significance testing. *Theory & Psychology* 18: 69-88.
- Hurlbert, S. H. & C. M. Lombardi. 2009. Final collapse of the Neyman-Pearson decision theoretic framework and rise of the neoFisherian. *Annales Zoologici Fennici* 46: 311-349.
- Johnson, D. H. 1999. The insignificance of statistical significance testing. *Journal of Wildlife Management* 63: 763-772.

- Keele, L. 2008. Semiparametric regression for the social sciences. Wiley, Hoboken. 213 p.
- Lindenmayer, D. B. & G. E. Likens. 2011. Losing the culture of ecology. *Bulletin of the Ecological Society of America* 92: 245-246.
- Lombardi, C. M. & S. H. Hurlbert. 2009. Misprescription and misuse of one-tailed tests. *Austral Ecology* 34: 447-468.
- MacKenzie, D. I., J. J. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey & J. E. Hines. 2006. *Occupancy estimation and modeling*. Elsevier, Boston. 324 p.
- Mielke, P. W. & K. J. Berry. 2001. *Permutation methods: a distance function approach*. Springer Verlag, Nueva York. 352 p.
- Nakagawa, S. & I. C. Cuthill. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews* 82: 591-605.
- Rodríguez, S. 1840. *Sociedades americanas en 1828. Primera parte. Luces y virtudes sociales*. Imprenta del Mercurio, Valparaíso. 82 p.
- Silva-Ayçaguer, L. C., P. Suárez-Gil & A. Fernández-Somoano. 2010. The null hypothesis significance test in health sciences research (1995-2006): statistical analysis and interpretation. *BMC Medical Research Methodology* 2010 (10) : <http://www.biomedcentral.com/1471-2288/10/44>.
- Stephens, P. A., S. W. Buskirk & C. M. del Río. 2007. Inference in ecology and evolution. *Trends in Ecology and Evolution* 22: 192-197.
- Váldez Galárraga, R. 2004. *Diccionario del pensamiento martiano*. Editorial de Ciencias Sociales, La Habana. 785 p.