

Conceptos y técnicas en ecología fluvial

Edición a cargo de:

ARTURO ELOSEGI

Profesor titular de Ecología en la Universidad del País Vasco

SERGI SABATER

Catedrático de Ecología en la Universidad de Girona

Separata del capítulo 22

Diseño experimental y análisis de datos

EMILI GARCÍA-BERTHOU

CARLOS ALCARAZ

LLUÍS BENEJAM

JOSEP BENITO

Primera edición: abril 2009

ISBN: 978-84-96515-87-1

© los autores, 2009

© de la edición en español, Fundación BBVA, 2009

Diseño experimental y análisis de datos

EMILI GARCÍA-BERTHOU, CARLOS ALCARAZ, LLUÍS BENEJAM Y JOSEP BENITO

«All models are wrong, but some are useful»
(BOX 1979)

«Normality is a myth; there never was, and never will be, a normal distribution»
(GEARY 1947)

22.1. Introducción

El diseño del estudio y el análisis de datos son partes cruciales en cualquier investigación ecológica. A menudo, el uso apropiado de diseños y análisis estadísticos es desafiante y frustrante para el investigador. Un primer error, todavía frecuente debido en parte a la aversión estadística de muchos biólogos, es pensar en la estadística sólo después de obtener los datos (Green 1979). Ello, a menudo, conduce a diseños inapropiados o subóptimos, dificultades de análisis y resultados dudosos.

Los estudios pueden dividirse en experimentos o *estudios experimentales* y *estudios observacionales*. En los experimentos, el investigador somete el material de estudio (unidad experimental) a diferentes tratamientos; por ejemplo, podemos fertilizar un tramo de río y compararlo con un control no fertilizado. En los estudios observacionales sólo observamos unidades muestrales en diferentes condiciones preestablecidas y no podemos decidir qué nivel o «tratamiento» recibirá una cierta unidad muestral; por ejemplo, podemos comparar las comunidades de diatomeas o macroinvertebrados a lo largo de un río o en lugares con distintos grados de contaminación. Las ventajas respectivas de los dos tipos de estudios pueden encontrarse, por ejemplo, en Hurlbert (1984) o Diamond (1986), pero pueden resumirse en un mayor control y claridad de resultados en los experimentos, y en mayores escalas espaciales y temporales, y mayor realismo de los estudios observacionales.

El pensamiento estadístico debe guiar el diseño experimental desde el principio

Tres principios importantes del diseño experimental: control, replicación y aleatorización

Al diseño de experimentos se ha prestado mucha atención desde los trabajos del gran estadístico y biólogo Ronald A. Fisher (Box 1978). El diseño de estudios observacionales se llama *diseño muestral* (*sampling design* en inglés), y desafortunadamente su aplicación ha sido mucho menos desarrollada en ecología que el diseño experimental. Tres principios clásicos del diseño experimental básicamente desarrollados por Fisher (véase, por ejemplo, Hurlbert 1984) son el control, la replicación (uso de varias unidades experimentales que reciben el mismo tratamiento) y la aleatorización (asignación al azar de los tratamientos a las distintas unidades experimentales).

El *control* es un grupo de unidades experimentales normalmente «no tratadas», con las que se compararán los tratamientos experimentales. El control proporciona un nivel de referencia que definirá los efectos de los tratamientos. Es importante ser cuidadosos con la definición del control, de modo que la única diferencia entre grupo control y tratamiento sea el factor que nosotros queremos estudiar: por ejemplo, los médicos usan los *placebos* (pastillas o preparados sin principios activos) y dan las pastillas *doblemente a ciegas* (es decir, ni el administrador ni el paciente saben si se está administrando un placebo o la sustancia que se está ensayando), porque el objetivo es sólo estudiar los efectos farmacológicos y no la autosugestión. Un equivalente ecológico sería que si queremos usar, por ejemplo, una *estructura de exclusión* para estudiar cuál es el efecto de los peces en la estructura trófica de un río, debemos evitar que la misma altere mucho las características del hábitat o fisicoquímicas (por ejemplo, aumentando la colmatación con sedimentos finos o reduciendo la luz), ya que de otro modo confundiremos fuentes de variación (depredación frente a estructura del hábitat).

Hay que tener cuidado en incurrir en la seudoreplicación

La *replicación* es el principio de que necesitamos dos o más unidades experimentales (réplicas) con el mismo tratamiento para medir el error experimental. La mayoría de métodos estadísticos (como el análisis de varianza o ANOVA) estudian la existencia de efectos de los tratamientos comparando la variación entre tratamientos con la variación dentro de los tratamientos. Aumentar el número de réplicas (y reducir el error experimental) aumentará la precisión de las estimaciones y la potencia de las pruebas estadísticas, es decir, la capacidad de detectar efectos. Un error frecuente en experimentos ecológicos es la *seudoreplicación*, término acuñado por Hurlbert (1984), que consiste en el análisis incorrecto que usa la variación dentro de la unidad experimental como error experimental. Por ejemplo, si sólo usamos un acuario tratado (por ejemplo, con un tóxico) y otro sin tratar, no podemos usar medidas de distintos individuos (por ejemplo, tamaño o fecundidad de invertebrados) dentro de cada acuario para analizar el efecto del tratamiento. Nótese que la seudoreplicación es un error más de análisis que de diseño, y que algunos experimentos sin replicación pueden analizarse correctamente, por ejemplo, mediante técnicas de regresión (Hurlbert 2004).

La *aleatorización* consiste en asignar al azar los tratamientos (y otros procedimientos) a las distintas unidades experimentales, para evitar sesgos del experimentador y favorecer la independencia de los errores. Aunque la aleatorización es, en principio, recomendable, de hecho seguramente es más importante la *interspersión* es decir, que los distintos tratamientos estén repartidos entre las distintas réplicas, evitando la agregación espacial de los distintos tratamientos (Hurlbert 2004).

Estos tres principios del diseño experimental son de fácil aplicación en el laboratorio pero, a menudo, presentan muchas dificultades para la experimentación en el campo a escalas medias o grandes, o para estudios como los de impactos ambientales (Carpenter et al. 1989, Carpenter 1990, Osenberg et al. 2006). La unidireccionalidad y heterogeneidad espacial de los ríos seguramente dificulta más el diseño experimental que en ecosistemas terrestres o lacustres (Downes et al. 2002). Imaginemos, por ejemplo, que queremos ver cuál es el efecto de fertilizar un tramo de río en los procesos biogeoquímicos o en las comunidades de diatomeas. Podemos añadir nutrientes en el tramo experimental y usar un control justo aguas arriba del tramo. Podríamos repetir simultáneamente la fertilización en otro tramo, aunque ello duplica el esfuerzo de campo. ¿Podemos usar la *variación* entre diferentes controles, por encima de un único tramo fertilizado? Pues, según Stewart-Oaten y Bence (2001), ello no sería correcto. Usar la variación entre un tramo fertilizado y dos controles por encima es inapropiado porque:

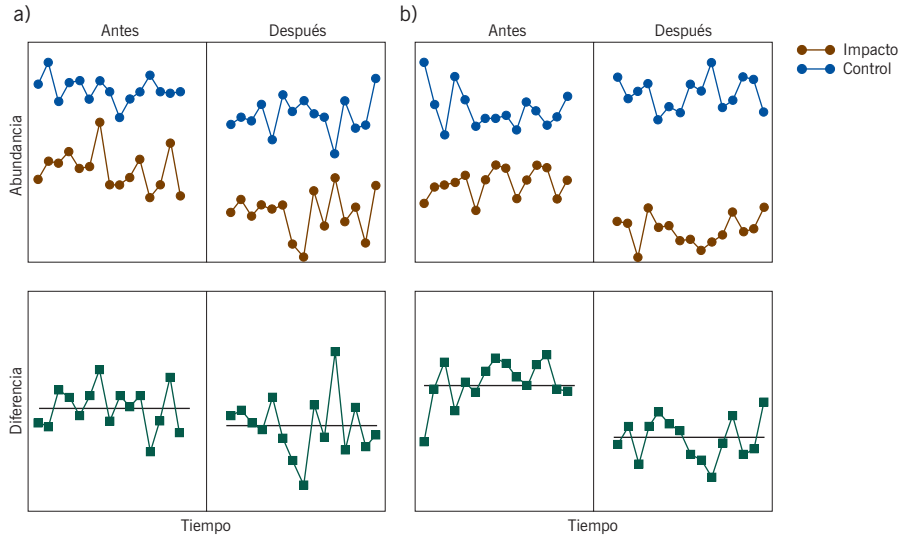
- estaremos confundiendo la variación entre tratamiento y control con la variación espacial entre controles;
- las unidades experimentales no son independientes y la variación entre ellas estará muy correlacionada y será función de la distancia entre ellas; y
- si usamos varios controles, debemos calcular la media entre ellos y comparar la trayectoria de esta media con la del tramo fertilizado.

En este caso, si no se pueden fertilizar varios tramos, el diseño más apropiado según Stewart-Oaten y Bence (2001) es un diseño BACI (*Before-After-Control-Impact*) también llamado BACIPS (*BACI-Paired Series*). El *diseño BACI* compara las series temporales (o las diferencias) de la zona tratada (o impactada) y el control, antes y después de la intervención o tratamiento experimental (fig. 22.1). Se pueden analizar con pruebas *t* de Student de las diferencias de medias o, preferiblemente, con técnicas de regresión o series temporales (si existen datos suficientes) (Stewart-Oaten y Bence 2001).

El diseño BACI consiste en comparar las condiciones en dos tramos, uno control y otro impactado, antes y después del impacto

A continuación proponemos algunos pasos generales para el diseño de estudios y análisis de datos ecológicos. Nuestro objetivo es compensar la desproporción entre matemáticas y aplicación de muchos libros y cursos estadísticos, a pesar de que en estadística tampoco son recomendables los *libros de recetas*, porque lo fun-

Figura 22.1:
Ilustración de un diseño BACI



Nota: Se estudia la abundancia de un organismo en una zona control y otra zona «impacto» antes y después de un determinado suceso, intervención, o tratamiento experimental. En la *izquierda* (a) existe una diferencia entre las dos zonas (control e impacto) y una tendencia temporal, pero que es similar en el control y la zona impactada o tratada. En cambio, en la *derecha* (b), las diferencias entre control e impacto cambian con la intervención (antes-después), por lo que concluiremos que existe un efecto de la intervención.

Fuente: Adaptada de Stewart-Oaten et al. (1992).

damental es entender bien los métodos para poder aplicarlos correctamente, y ningún libro incluye todos los casos posibles (Box 1976, Hurlbert 1984).

Técnica 55. Diseño del estudio

ALGUNOS PRINCIPIOS GENERALES (O LOS DIEZ MANDAMIENTOS DE GREEN)

Los diez mandamientos de Green son un compendio de buenas conductas en estadística aplicada

A modo de principios generales, enumeramos, a continuación, una adaptación de lo que nosotros llamamos los *diez mandamientos* de Green (1979), advirtiendo que «quien esté libre de pecado, que tire la primera piedra»:

1. *Objetivo.* Serás capaz de explicar a cualquier persona, de forma clara y concisa, qué cuestión quieres resolver. Los resultados del estudio serán tan coherentes y comprensibles como tu concepción inicial del problema.
2. *Replicación.* Tomarás réplicas por cada combinación de tiempo, lugar y cualquier otro factor controlado. Las diferencias *entre* grupos sólo pueden detectarse comparándolas con las diferencias *dentro* de estos grupos.
3. *Aleatorización.* Tomarás aleatoriamente *un número igual* de réplicas para cada combinación de variables controladas. Decidir qué réplicas son «representativas» o «típicas» no es un muestreo aleatorio.

4. *Control*. Con el fin de comprobar si una condición produce un efecto, es necesario muestrear en presencia y en ausencia de la condición, sin que el resto de condiciones varíen. Un efecto sólo se puede demostrar por comparación con un control.
5. *Estudio piloto*. Tomarás muestras preliminares con el fin de obtener información básica que permita evaluar qué diseño muestral y qué análisis estadístico son los más adecuados. Quien se salta este paso por falta de tiempo, normalmente acaba perdiendo más tiempo.
6. *Eficiencia del muestreo*. Verificarás que el método o aparato de muestreo está realmente muestreando la población que piensas, con eficiencia constante y adecuada en todo el rango de condiciones de muestreo. Las diferencias en la eficiencia de muestreo sesgan las comparaciones.
7. *Muestreo estratificado, cuando hay heterogeneidad poblacional*. Si la población muestra heterogeneidad es necesario realizar un muestreo aleatorio estratificado: dividirás la población en estratos homogéneos y tomarás muestras proporcionalmente según el tamaño del estrato. Por ejemplo, si las comunidades que estudias parecen diferir entre rápidos y pozas, muestrearás cada uno de estos hábitats por separado.
8. *Unidad experimental o muestral, y número de réplicas (tamaño muestral)*. Definirás apropiadamente la unidad experimental o muestral. Estimarás el número mínimo de réplicas necesario para obtener la precisión deseada.
9. *Verificarás los supuestos de los análisis estadísticos*. Comprobarás si las varianzas son homogéneas y si los errores tienen una distribución normal. Si, como es habitual en biología, esto no se cumple: a) transformarás apropiadamente los datos, b) utilizarás un procedimiento no paramétrico, c) utilizarás un diseño muestral secuencial apropiado, o d) usarás métodos de aleatorización o permutación.
10. *Escogerás el método estadístico al diseñar el estudio*. Una vez escogido el método estadístico para comprobar la hipótesis, aceptarás el resultado. Un resultado no esperado o no deseado no es una razón válida por rechazar el método y buscar uno «mejor».

Estos principios son más esenciales en experimentos de menor tamaño. En estudios observacionales o experimentos a gran escala, el diseño es más complicado y se permiten más «pecados». Por ejemplo, la replicación es lógicamente muy recomendable, pero las técnicas de regresión son aplicables correctamente sin replicación.

ESTRUCTURA DEL DISEÑO EXPERIMENTAL O MUESTRAL

Un primer paso para diseñar bien un estudio y, por ende, escoger el método estadístico adecuado, es aplicar bien la terminología del diseño, que es un poco distinta para experimentos y estudios observacionales, a la estructura de nuestro estudio (cuadro 22.1).

Cuadro 22.1:
Terminología del diseño experimental (experimentos) y muestral (estudios observacionales)

Definición	Estudio observacional	Experimento
Aquellas <i>variables</i> que nosotros manipulamos (experimento) o que medimos a niveles distintos ya preestablecidos (estudio observacional)	Factor o variable independiente	Factor, variable experimental o variable independiente
Variable que medimos y queremos saber cómo responde a nuestra manipulación o a los distintos niveles del factor	Variable, variable de respuesta o variable dependiente	Variable, variable de respuesta o variable dependiente
Distintos niveles de un factor	Nivel	Tratamiento o nivel
Todo el conjunto de material al cuál aplicamos un cierto tratamiento, independientemente del resto de material. Las distintas unidades experimentales que reciben el mismo tratamiento reciben el nombre de réplicas	–	Unidad experimental
Conjunto de material de un cierto nivel del factor, del cual medimos una o más variables. A menudo los ecólogos las llamamos muestras	Unidad muestral	–
Conjunto de unidades muestrales con los mismos niveles del factor o factores (concepto tradicional de muestra en estadística)	Grupo o muestra	Grupo, muestra o combinación de tratamientos

Fuente: Modificado de Sachs (1978).

Los estudios observacionales y los experimentos siguen aproximaciones estadísticas diferentes, lo que se refleja en la terminología utilizada

Si aplicamos esta terminología a nuestros estudios, podremos decidir qué métodos estadísticos utilizar. Por ejemplo, el de la figura 22.1 será un estudio observacional si se trata de un impacto ambiental que ya se ha producido, y será un experimento si es una fertilización o adición de nutrientes controlada por el investigador. En cualquier caso, existen dos factores: la *zona* (con los niveles control e impacto) y el *tiempo* (que tiene múltiples niveles pero del que, básicamente, queremos decidir si existe una diferencia antes y después de la intervención). La *variable de respuesta* es la abundancia de un cierto organismo o cualquier variable que nosotros queramos estudiar cómo se ve afectada por el impacto. La *unidad muestral* o *experimental* es el material en cada una de las zonas en un momento determinado de medición. Podemos considerar que tenemos cuatro muestras resultantes de dos factores con dos niveles cada uno, pero sin replicación. En selección de métodos de la técnica 56 veremos cómo analizar este tipo de datos.

Técnica 56. Análisis de datos

SELECCIÓN DE PROGRAMA INFORMÁTICO

Existen actualmente varios paquetes estadísticos muy completos (por ejemplo, SAS, SPSS, S-plus). La elección entre ellos, a menudo, depende de la disponibili-

dad y tradición de uso en nuestro entorno más inmediato. Sí que es poco recomendable ceñirse sólo a Excel, ya que sus extensiones son limitadas y constriñen la elección de métodos.

Es esencial manejar programas estadísticos completos

Una auténtica revolución estadística es actualmente el entorno R que, a pesar de ser más difícil de aprender, es gratuito y tiene las siguientes ventajas:¹

- Tiene revisiones y actualizaciones muy frecuentes, y constante aparición de paquetes con nuevos métodos; los investigadores estadísticos que desarrollan actualmente nuevos métodos los implementan a menudo en R, por lo que se facilita mucho la transición entre desarrollo de nuevos métodos y su aplicación generalizada.
- Están apareciendo versiones Gui (entorno gráfico similar a Windows o a Macintosh) de algunas aplicaciones, por lo que se facilita el uso.
- Existen numerosos métodos específicamente desarrollados o recomendados para datos ecológicos.² Desarrollado por ecólogos de la Universidad Claude-Bernard Lyon I originalmente para Macintosh, incluye muchas más variantes de técnicas de ordenación (Dolédec et al. 2000, Pelissier et al. 2003, Dray et al. 2003) que las que pueden hacerse con un *software* de más fácil uso pero comercial como CANOCO (Ter Braak y Smilauer 2002, Lepš y Smilauer 2003) que, seguramente, aún es el más utilizado en ecología.
- Flexibilidad para implementar métodos o cálculos propios (por ejemplo, simulaciones).
- Facilidad de repetición de los mismos análisis (ya que se genera un *script*) por el mismo u otros usuarios.

El entorno R ofrece gran variedad de aplicaciones estadísticas gratuitas, aunque no es el programa de más fácil manejo

ANÁLISIS EXPLORATORIO DE DATOS

En cualquier paso de la investigación científica se cometen errores. Una vez introducidos los datos en el programa estadístico elegido, es muy recomendable realizar lo que se llama *análisis exploratorio de datos*, el cual, además de servir para depurar los datos, también sirve para entender las características de los mismos (por ejemplo, la forma de la distribución de las variables de respuesta). Este análisis exploratorio de datos puede consistir en estadística descriptiva, histogramas, gráficas de caja, etc. Resultan muy útiles las gráficas de dispersión entre variables cuantitativas, ya que a menudo nos permiten identificar valores atípicos o *outliers*. Debemos revisar estos valores atípicos y eliminarlos tan sólo en aquellos casos

Un análisis exploratorio de los datos es fundamental para depurar los mismos y comprender el tipo de distribución de las variables

¹ Puede encontrarse información sobre este entorno en <http://www.r-project.org/>

² Véase una relación de ellos en la *Task View: Analysis of Ecological and Environmental Data*, disponible en <http://cran.r-project.org/web/views/Environmetrics.html>. Por ejemplo, el programa ade4, consultable en <http://pbil.univ-lyon1.fr/ADE-4/>

en que estemos convencidos de que se trata de un dato altamente improbable, fruto de error de lectura o anotación; lo mejor es no eliminarlos dada la variabilidad del fenómeno.

SELECCIÓN DE MÉTODOS

Para seleccionar los métodos estadísticos, previa aplicación de la terminología del diseño (véase técnica anterior), debemos establecer las características siguientes para nuestros datos (entre paréntesis la técnica a utilizar):

A fin de seleccionar la técnica estadística adecuada cabe responder estas siete cuestiones

1. *Tipo de estudio.* Experimento con variables cuantitativas (contraste de hipótesis o estadística inferencial), estudio observacional de ecología de comunidades (técnicas de ordenación y clasificación de análisis multivariante).
2. *Tipo de variable/nivel de medida.* Cuantitativa o medida (estadística paramétrica como el análisis de la varianza, la regresión múltiple o los modelos lineales generalizados), ordinal (estadística no paramétrica), categórica binaria o politémica (análisis de frecuencias o de datos categóricos: regresión logística o modelos lineales generalizados con errores binomiales, etc.).
3. *Número de variables.* Una (estadística univariante) o dos variables de respuesta, o bien muchas variables en general (análisis multivariante).
4. *Número de factores.* Uno o más; número de muestras: uno, dos o más.
5. *Objetivo/hipótesis.* Estudiar la tendencia central, la variabilidad, toda la forma de la distribución (pruebas de bondad de ajuste), la relación entre variables (correlación, regresión, análisis multivariante).
6. *¿Tenemos muchos datos ($n > 50$) durante el tiempo (series temporales), o el espacio (estadística espacial)?*
7. *¿Tenemos datos de supervivencia? (análisis de supervivencia)*

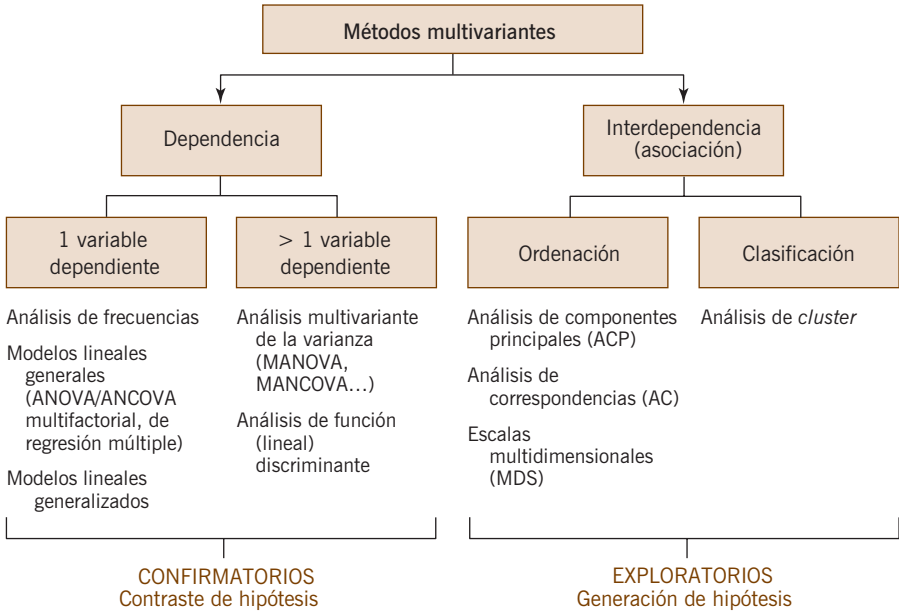
Si respondemos a estas cuestiones podremos escoger entre las principales técnicas univariantes mediante el cuadro 22.1 y entre las multivariantes mediante la figura 22.2. Volvamos a nuestro ejemplo de la figura 22.1 para intentar aplicar estos siete puntos. Imaginemos que medimos para el diseño de la figura 22.1, unas pocas variables de respuesta cuantitativas (por ejemplo, abundancia de unas pocas especies o algunas variables fisicoquímicas). La respuesta a estos siete puntos sería:

1. Si medimos unas pocas variables de respuesta cuantitativas, el objetivo será comparar el control con la zona impactada, y necesitamos estadística inferencial (es decir, contraste de hipótesis).
2. Entre la estadística inferencial, como las variables son cuantitativas usaremos estadística paramétrica. Si tenemos variables fisicoquímicas o abundancias elevadas con pocos cerros, podemos usar modelos lineales (ANOVA, regresión). Si tenemos variables con unos pocos valores discretos (abundancias bajas con

<p>Tipos de variable ESTADÍSTICA ↓</p>	<p>2 o más muestras dependientes; 1 var. y más que 1 factor cuantitativo o 1 factor con medidas repetidas</p> <p>2 o más muestras independientes; 1 var. y 1 factor</p> <p>1 muestra</p>	<p>2 o más muestras dependientes; 1 var. y más que 1 factor cuantitativo o 1 factor con medidas repetidas</p> <p>2 o más muestras independientes; 1 var. y 1 factor</p>	<p>Prueba de ajuste a proporción teórica Prueba de independencia Regresión logística Modelos lineales generalizados</p> <p>Spearman (r_s) Kendall (τ)</p> <p>Mann-Whitney</p> <p>Runs</p> <p>Kolmogorov-Smirnov Shapiro-Wilks</p> <p>Prueba χ^2</p> <p>Prueba t</p>	<p>Objetivo ↓</p>
<p>Nominales o categóricas ANÁLISIS DE FRECUENCIAS</p>	<p>Relación</p>	<p>Relación</p>	<p>Relación</p>	<p>Objetivo ↓</p>
<p>Ordinales ESTADÍSTICA NO PARAMÉTRICA</p>	<p>Relación</p>	<p>Tendencia central</p>	<p>Relación</p>	<p>Objetivo ↓</p>
<p>Cuantitativas ESTADÍSTICA PARAMÉTRICA</p>	<p>Forma</p>	<p>Forma</p>	<p>Forma</p>	<p>Objetivo ↓</p>

Cuadro 22.2:
Selección de métodos univariantes en función de las características del diseño del estudio

Figura 22.2:
 Selección de métodos multivariantes en función del tipo de datos y el objetivo del estudio



ceros), será preferible usar modelos lineales generalizados (con errores de Poisson para las abundancias).

3. Si tenemos una o muy pocas variables, usaremos estadística univariable. Si tenemos muchas variables, a menudo es útil un análisis multivariante, por ejemplo, el análisis de varianza multivariante MANOVA (véase, por ejemplo, García-Berthou y Moreno-Amich, 1993) o el análisis discriminante; método que está muy relacionado con el anterior.
4. Ya hemos comentado que puede considerarse que tenemos cuatro muestras resultantes de dos factores (antes-después y control-impacto) con dos niveles cada uno.
5. El objetivo en ecología es, en general, estudiar las medias, pero a menudo también sería interesante estudiar las varianzas, aunque metodológicamente es más complicado. En los diseños BACI, en general, se usan el test *t* o ANOVA (cuyas hipótesis hacen referencia a la tendencia central), aunque las técnicas de regresión seguramente son más flexibles.
6. Si tenemos series largas de tiempo ($n > 50$) usar técnicas de series temporales es muy recomendable en un diseño BACI, aunque en ecología no es frecuente tener datos suficientes.
7. Existen métodos desarrollados para tipos de datos específicos, por ejemplo, el análisis de supervivencia cuando nuestra variable de respuesta es la supervivencia de individuos (se trata de un tipo de datos denominado *censurado* porque algunos individuos que han sobrevivido durante todo el experimento pueden morir muy poco después).

En cambio, si para el mismo diseño hemos estudiado la composición de algas o macroinvertebrados (con decenas de especies) y queremos estudiar cómo varía la composición de especies, la respuesta básicamente diferirá en los tres primeros puntos:

1. Se trata de un estudio de ecología de comunidades, donde típicamente tenemos decenas de variables (abundancia de muchas especies) y la estadística inferencial es menos apropiada (aunque las técnicas de ordenación también admiten algún diseño BACI (Lepš y Smilauer 2003)). Por ello, usaremos métodos como las técnicas de ordenación o clasificación (como, por ejemplo, el análisis de correspondencias), que en lugar de métodos de estadística inferencial y confirmación de hipótesis, típicamente son métodos exploratorios y generadores de hipótesis (fig. 22.2).
2. Si tenemos unas pocas variables de presencia o abundancia de especies, éstas son las variables de respuesta (cuantitativas), que se podrían analizar con modelos lineales generalizados, o para datos de presencia-ausencia, con regresión logística. Pero si tenemos decenas de especies y el objetivo es comparar la composición de especies, entonces la variable de respuesta mejor es la «especie», que es categórica *politómica*, y de las que hemos medido las frecuencias u otra estima de abundancia; por ello, habitualmente se usa el *análisis de correspondencias* (que se basa en la descomposición de la tabla de contingencia y la distancia chi-cuadrado).
3. Las técnicas de ordenación y clasificación, tan frecuentemente usadas en ecología de comunidades, se consideran técnicas de análisis multivariante, ya que analizan no unas pocas variables o factores, sino grandes tablas (de especies \times muestras y variables ambientales \times muestras) de muchas *variables*.

Entre los métodos de ordenación (fig. 22.2), podemos escoger entre las técnicas siguientes según el objetivo y tipo de datos (Ter Braak 1988):

1. Análisis indirecto de gradiente.
 - 1.1. Métodos basados en distancias
 - Análisis de coordenadas principales (*metric multidimensional scaling*).
 - Escalas multidimensionales no métricas o *nonmetric multidimensional scaling* (NMDS).
 - 1.2. Álgebra matricial (*eigenanalysis*).
 - 1.2.1. Respuesta lineal.
 - Análisis de componentes principales (ACP).
 - 1.2.2. Respuesta unimodal
 - Análisis de correspondencias (AC) o *reciprocal averaging*.
 - Análisis de correspondencias sin tendencia o *detrended correspondence analysis* (DCA).

2. Análisis directo de gradiente.
 - 2.1. Respuesta lineal
 - Análisis de redundancia (RDA).
 - 2.2. Respuesta unimodal
 - Análisis canónico de correspondencias (CCA).

Las técnicas de análisis indirecto analizan una sola matriz de datos, las de análisis directo se basan en dos matrices

El *NMDS* es un método basado en distancias ampliamente usado en psicología y también en biología marina.³ En cambio, en ecología terrestre y limnología, los métodos más utilizados son las distintas variedades de ACP y análisis de correspondencias.

Las técnicas de *análisis indirecto de gradiente* analizan una sola matriz, por ejemplo, un AC para la abundancia de diatomeas en distintas muestras o un ACP para la fisicoquímica. En cambio, las técnicas de *análisis directo de gradiente* o *análisis restringido* (*constrained*) como el CCA, analizan al mismo tiempo dos matrices, construyendo que la ordenación de las muestras, según la composición de especies, sea una función lineal (modelo de regresión múltiple) de las variables ambientales (físicoquímicas). Ambos tipos de técnicas (indirectas y directas o *constrained*) tienen objetivos distintos y son alternativas o complementarias: aunque el CCA resume en un solo análisis toda la información, es posible que la principal variación en la composición de especies corresponda a una variable ambiental no medida, por lo que un AC mostraría una ordenación muy distinta (y muy informativa) (Økland 1996). A menudo el CCA es menos robusto y precisa muchas más muestras que variables ambientales (McCune 1997). Buenos manuales para el uso de estas técnicas son Jongman et al. (1995) y Lepš y Smilauer (2003).

VERIFICACIÓN DE SUPUESTOS DEL MODELO

Es importante verificar los supuestos del método estadístico utilizado, pero sin caer en extremismos

La verificación de los supuestos o condiciones en los que se basan los métodos estadísticos es una de las cuestiones que más preocupan a los biólogos, y en las que más malentendidos persisten (Box 1953, Stewart-Oaten 1995, Smith 1995, Johnson 1995). A continuación intentamos aclarar algunos:

1. Las variables biológicas, en general, no siguen *exactamente* la distribución normal, sino más bien la lognormal. La *distribución lognormal* es similar a la normal (ambas son unimodales, es decir, con un solo máximo), pero no puede tomar valores negativos y tiene asimetría positiva. Por eso, las variables biológicas, en general, tienen una distribución *aproximadamente* normal pero la transformación logarítmica (de cualquier tamaño, concentración de sustancias...), suele

³ Véase *software* Primer en <http://www.zen87707.zen.co.uk/primer-e/>

mejorar el ajuste a la normalidad. Por ejemplo, si una variable siguiera la distribución normal tendría que existir entre menos y más infinito. Pero eso raramente se cumple en biología: los tamaños, pesos, concentraciones, etc., no pueden tomar valores negativos.

2. Las variables biológicas suelen mostrar *heteroscedasticidad*, es decir, cuando las medias son diferentes las varianzas también. A menudo, las desviaciones típicas aumentan linealmente con la media y, por eso, la transformación logarítmica tiende a arreglar también el supuesto de homoscedasticidad u homogeneidad de varianzas.
3. La estadística paramétrica es muy *robusta* (poco sensible) a la no normalidad; es decir, sus resultados tienden a ser correctos aunque no se cumpla el supuesto de normalidad. El supuesto de normalidad es poco relevante para la estadística paramétrica.
4. La estadística paramétrica es robusta a la heteroscedasticidad, sobre todo si la replicación es constante (Box 1953). Por lo tanto, el supuesto de homoscedasticidad es más importante que el de normalidad.
5. La robustez del análisis de la varianza (ANOVA) se manifiesta en que, en general, su conclusión es la misma tanto si se considera la variable original como la transformación logarítmica. Sin embargo, en un pequeño porcentaje de casos, la conclusión con la transformación es diferente y más fiable, por lo que es importante transformar los datos apropiadamente o bien usar modelos lineales generalizados (con la distribución de errores adecuada).
6. Las pruebas de homoscedasticidad (sobre todo la de Bartlett) son poco robustas a la no normalidad. Es decir, paradójicamente verificamos supuestos poco importantes para el ANOVA con métodos en que el supuesto de normalidad sí es importante.
7. En el análisis de regresión (y correlación) lineal, el supuesto de linealidad es más importante que los de homoscedasticidad y normalidad, ya que si la relación es claramente no lineal los estadísticos (r , recta de regresión) describirán muy mal los datos. El sesgo de los estadísticos generalmente será mucho mayor si hay no linealidad que si hay heteroscedasticidad o no normalidad.
8. A medida que aumenta el número de datos (réplicas), los resultados de todos los tests son más fiables (propiedad fundamental denominada *consistencia estadística*). Los tests de normalidad tienden a ser poco potentes (muy conservadores), es decir, a no detectar pequeñas desviaciones de la normalidad (resultados no significativos incorrectos, o lo que es lo mismo, errores de tipo II). Cuando aumenta la replicación es más fácil detectar estas pequeñas desviaciones y, por lo tanto, tendemos a encontrar mayor no normalidad; pero cuando aumenta la replicación, el supuesto de normalidad afecta menos al ANOVA. Otra paradoja: detectamos más el no cumplimiento de la normalidad cuando es menos importante.
9. Los métodos gráficos para estudiar la homoscedasticidad y la normalidad son más subjetivos y menos claros que los tests. Pero de hecho, por las paradojas

Al aumentar el número de réplicas, aumenta la fiabilidad de los tests estadísticos

de los puntos 6 y 8, los métodos gráficos son más fiables y útiles (en contra de lo que parece inicialmente), porque nos indican si es adecuado utilizar una transformación y cuál de ellas.

10. A pesar de lo que se piensa muy a menudo, la estadística no paramétrica también hace supuestos. Por ejemplo, el test de Mann-Whitney y el de Kruskal-Wallis suponen que la única diferencia, si la hay, entre los grupos que se comparan es en la tendencia central; y que la dispersión y distribución de los grupos es igual excepto con respecto a la tendencia central (Johnson 1995, Kasuya 2001). Es decir, hacen supuestos similares a la homoscedasticidad y no son ninguna panacea como se suele asumir. En cambio, para las pruebas *t* de Student hay fórmulas que no suponen la homoscedasticidad, por lo que realizan menos supuestos y sus resultados a menudo serán más correctos que los de un método no paramétrico (Johnson 1995). Lo mismo ocurre con los tests de aleatorización o permutación: hacen los mismos supuestos que los paramétricos, excepto el de normalidad (que es irrelevante para estos últimos) (Stewart-Oaten et al. 1992, Hayes 2000).
11. Sólo existen los equivalentes no paramétricos de los diseños experimentales más sencillos. Por ejemplo, la mayoría de paquetes estadísticos no tienen el equivalente no paramétrico de un diseño de bloques aleatorizados o de un ANOVA de medidas repetidas. Por ello, a menudo, se aplican métodos no paramétricos de diseños más simples sin respetar el diseño experimental establecido, lo que puede conducir fácilmente a conclusiones erróneas.

Es más recomendable transformar adecuadamente los datos y aplicar métodos paramétricos que usar métodos no paramétricos

En resumen, los métodos no paramétricos son «sirenas estadísticas» (Johnson 1995) poco recomendables (con la excepción del *coeficiente de correlación de Spearman*, que es una manera rápida de estudiar correlaciones sin buscar la transformación adecuada). En lugar de aplicar métodos no paramétricos, recomendamos estudiar gráficamente (en vez de con contrastes de hipótesis) la distribución de nuestras variables y usar una de las dos opciones siguientes:

1. Aplicar una transformación a los datos y usar métodos paramétricos como los modelos lineales (ANOVA o regresión lineal) o las pruebas *t* de Student. Las transformaciones principales a usar (que son simplemente un cambio de escala que puede tener mucho sentido) son: la logarítmica (para variables medidas como concentraciones, pesos, tamaños, ...), la raíz cuadrada (para variables también con asimetría positiva como frecuencias) y el arcoseno ($\arcsin \sqrt{x}$, para las proporciones) (Sokal y Rohlf 1995).
2. Aplicar *modelos lineales generalizados* (GLM), que extienden los modelos lineales a muchas otras distribuciones y permiten ajustar los supuestos de tipo de distribución, homoscedasticidad y linealidad, de forma independiente. Los GLM (McCullagh y Nelder 1989) incluyen los modelos lineales como caso particular (cuando se usan errores normales y función *link* identidad) y podemos

utilizar errores de Poisson para variables de frecuencia (por ejemplo, número de individuos) y errores binomiales para las proporciones o variables de presencia-ausencia (o incluso la regresión logística, que es matemáticamente equivalente a un GLM con errores binomiales y función *link* logit).

22.2. Bibliografía

- BOX G.E.P. «Non-normality and tests on variances». *Biometrika* 40 (1953): 318-335.
- BOX G.E.P. «Science and statistics». *Journal of the American Statistical Association* 71 (1976): 791-799.
- BOX G.E.P. «Robustness in the strategy of scientific model building». (Citado en Cook y Weisberg 1997.) En R.L. Launer, y G.N. Wilkinson, eds. *Robustness in statistics*. Nueva York: Academic Press, 1979: 201-236.
- BOX J.F. *R.A. Fisher. The life of a scientist*. Nueva York: Wiley, 1978.
- CARPENTER S.R. «Large-scale perturbations: Opportunities for innovation». *Ecology* 71 (1990): 2038-2043.
- CARPENTER S.R, FROST T.M, HEISEY D., y KRATZ T.K. «Randomized intervention analysis and the interpretation of whole-ecosystem experiments». *Ecology* 70 (1989): 1142-1152.
- COOK R.D., y WEISBERG S. «Graphics for assessing the adequacy of regression models». *Journal of the American Statistical Association* 92 (1997): 490-499.
- DIAMOND J. «Overview: Laboratory experiments, field experiments, and natural experiments». En J. Diamond, y T.J. Case, eds. *Community Ecology*. Nueva York: Harper and Row, 1986: 3-22.
- DOLÉDEC S., CHESEL D., y GIMARET-CARPENTIER C. «Niche separation in community analysis: A new method». *Ecology* 81 (2000): 2914-2927.
- DOWNES B.J., BARMUTA L.A., FAIRWEATHER P.G., FAITH D.P., KEOUGH M.J., LAKE P.S., MAPSTONE B.D., y QUINN G.P. *Monitoring ecological impacts: Concepts and practice in flowing waters*. Cambridge: Cambridge University Press, 2002.
- DRAY S., CHESEL D., y THIOULOUSE J. «Co-inertia analysis and the linking of ecological data tables». *Ecology* 84 (2003): 3078-3089.
- GARCIA-BERTHOU E., y MORENO-AMICH R. «Multivariate analysis of covariance in morphometric studies of the reproductive cycle». *Canadian Journal of the Fisheries and Aquatic Sciences* 50 (1993): 1394-1399.
- GEARY R.C. «Testing for normality». *Biometrika* 34 (1947): 209-242.
- GREEN R.H. *Sampling design and statistical methods for environmental biologists*. Nueva York: John Wiley and Sons, Inc., 1979.
- HAYES A.F. «Randomization tests and the equality of variance assumption when comparing group means». *Animal Behaviour* 59 (2000): 653-656.
- HURLBERT S.H. «Pseudoreplication and the design of ecological field experiments». *Ecological monographs* 54 (1984): 187-211.
- HURLBERT S.H. «On misinterpretations of pseudoreplication and related matters: A reply to Oksanen». *Oikos* 104 (2004): 591-597.
- JOHNSON D.H. «Statistical sirens: The allure of nonparametrics». *Ecology* 76 (1995): 1998-2000.
- JONGMAN R.H.G., TER BRAAK C.J.F., y VAN TONGEREN O.F.R., eds. *Data analysis in community and landscape ecology*. Cambridge: Cambridge University Press, 1995.

- KASUYA E. «Mann-Whitney U test when variances are unequal». *Animal Behaviour* 61 (2001): 1247-1249.
- LEPŠ J., y ŠMILAUER P. *Multivariate analysis of ecological data using CANOCO*. Cambridge: Cambridge University Press, 2003.
- MCCULLAGH P., y NELDER J.A. *Generalized linear models*. Nueva York: Chapman and Hall, 1989.
- MCCUNE B. «Influence of noisy environmental data on canonical correspondence analysis». *Ecology* 78 (1997): 2617-2623.
- ØKLAND R.H. «Are ordination and constrained ordination alternative or complementary strategies in general ecological studies?». *Journal of Vegetation Science* 7 (1996): 292.
- OSENBERG C.W., M. B.B., WHITE J.S., ST. MARY C., y SHIMA J.S. «Statistical issues and assessment of ecological restorations: Lessons from marine reserves». En D. Falk, N. Palmer, and J. Zedler, eds. *Foundations of restoration ecology*. Washington, D.C.: Island Press, 2006: 280-302.
- PELISSIER R., COUTERON P., DRAY S., y SABATIER D. «Consistency between ordination techniques and diversity measurements: Two strategies for species occurrence data». *Ecology* 84 (2003): 242-251.
- SACHS L. *Estadística aplicada*. Barcelona: Ed Labor, 1978.
- SMITH S.M. «Distribution-free and robust statistical methods: Viable alternatives to parametric statistics?». *Ecology* 76 (1995): 1997-1998.
- SOKAL R.R., y ROHLF F.J. *Biometry: The principles and practice of statistics in biological research*. Nueva York: Freeman, 1995.
- STEWART-OATEN A. «Rules and judgments in statistics: Three examples». *Ecology* 76 (1995): 2001-2009.
- STEWART-OATEN A., y BENICE J.R. «Temporal and spatial variation in environmental impact assessment». *Ecological monographs* 71 (2001): 305-339.
- STEWART-OATEN A., BENICE J.R., y OSENBURG C.W. «Assessing effects of unreplicated perturbations: No simple solutions». *Ecology* 73 (1992): 1396-1404.
- TER BRAAK C.J.F. «A theory of gradient analysis». *Advances in Ecological Research* 18 (1988): 271-317.
- TER BRAAK C.J.F., y ŠMILAUER P. *CANOCO reference manual and CANODRAW for Windows user's guide: Software for canonical community ordination (version 4.5)*. Ithaca, Nueva York: Micro-computer Power, 2002.