

PROGRAMA DE MÁSTER Y DOCTORADO: “Biodiversidad en Áreas Tropicales y su Conservación”

Asignatura: Diseño y metodologías de investigación en biología de la conservación.

OBJETIVOS:

- Conocimiento por el alumno de las técnicas metodológicas, más en concreto, las utilizadas en el ámbito experimental.
- Control, elaboración y aplicación correcta de dicha metodología según las características del problema a resolver.
- Desarrollo de su capacidad crítica y un mayor rigor científico.

La meta a alcanzar mediante estos objetivos es que el alumno mediante la adquisición de un método de trabajo (científico), pueda actuar con un criterio regular, sistemático y ordenado, llevando a cabo una planificación experimental que le permita resolver aquellos problemas científicos que se plantee.

La estrategia principal para conseguir un buen aprendizaje se basa en el análisis detallado (seguimiento, crítica, valoración...) de cada uno de los diseños experimentales desarrollados por cada alumno en presencia del resto. Las clases prácticas son, por tanto, participativas, donde tanto el profesor como los alumnos han de conseguir optimizar lo más posible los diferentes diseños expuestos. Las exposiciones se realizarán con apoyo del material audiovisual que el alumno precise (transparencias, diapositivas, cañón...).

PROGRAMACIÓN

- Clases teoría: 10 horas
- Clases prácticas: 30 horas.

Todas las horas lectivas son presenciales.

Las clases prácticas consisten en la elaboración por el alumno de un diseño experimental que permita cumplir de forma lógica y eficiente el objetivo general del trabajo por él elegido o, en su defecto, recomendado por el profesor. El desarrollo metodológico de este trabajo será seguido muy directamente por el profesor, realizándose siempre en presencia del resto de los alumnos, lo que permite optimizar el aprendizaje sin demasiado esfuerzo, además de ofrecer una amplia gama temática.

CRITERIOS DE EVALUACIÓN

El proceso de evaluación se basará en el nivel de participación, tanto en el desarrollo de su trabajo como en el de sus compañeros, así como la exposición oral y escrita (memoria).

TEMARIO

El método científico

- Tipos de Investigaciones: documental, de casos, observacional y experimental.
- El método científico. El método inductivo y el deductivo. Pasos en el método científico.

El diseño experimental

- Conceptos generales. Definición. Unidad experimental, unidad de muestreo, factores, tratamientos y variables.
- Etapas fundamentales de la experimentación:
 - Reconocimiento de que un problema existe.
 - Formulación del problema.
 - Establecimiento de factores y niveles.
 - Especificaciones de las variables.
 - Definición de la inferencia espacial del problema.
 - Selección al azar de las unidades experimentales.
 - Asignación de los tratamientos a las unidades experimentales.
 - Perfilar los análisis antes de tomar los datos.
 - Toma de datos.
 - Análisis de datos.
 - Conclusiones.
 - Implementación.
- Análisis detallado del diseño experimental:
 - Unidades de muestreo y variables. Tipos de variables.
 - Muestreo. Replicación. Exactitud y precisión.
 - Selección de las unidades de muestreo. Tipos de muestreo.
 - Pseudorreplicación. Tipos: simple, temporal y sacrificada.
 - Determinación del tamaño de muestra.
 - Estadísticos básicos. Medidas de tendencia general y medidas de variabilidad.
 - Contraste de hipótesis. Errores tipo I y II. Violación de la normalidad.
 - Aplicación de pruebas estadísticas sencillas. Pruebas de comparación de medias, de comparación de frecuencias y pruebas de asociación.
 - Consideraciones sobre el diseño experimental. Controles, replicación, aleatorización y entremezcla. Problemas y soluciones.
 - La importancia de la representatividad.
 - Errores en el diseño experimental: errores de diseño, aleatorios y sistemáticos.

BIBLIOGRAFÍA

- Anderson, V.L. & McLean, R.A. (1974). Design of experiments. A Realistic Approach. Marcel Dekker. New York. 418 pp.
- Cochran, W.G. (1980). Técnicas de muestreo. CECSA. México. 513 pp.
- Cochran, W.G. y Cox, G.M. (1991). Diseños experimentales. Impresora Roma. México. 661 pp.
- Cox, D.R. (1992). Planning of Experiments. John Wiley & Sons. New York. 308 pp.
- Elena Roselló, J.M. y Fernández de Gorostiza, M. (1986). Guía técnica para ensayos de variedades de campo. FAO. Roma.
- **Feinsinger, P. (2003). El diseño de estudios de campo para la conservación de la biodiversidad. Ed. FAN. Santa Cruz de la Sierra. Bolivia. 242 pp.**
- García Roldán, J.L. (1995). Cómo elaborar un proyecto de investigación. Universidad de Alicante. Alicante. 173 pp.
- Garton, E.O. & Ratti, J.T. (1996). Research and experimental design. En: Research and management techniques for wildlife and habitats. Bookhout, T.A. (ed.). The Wildlife Society. Bethesda, Maryland. 740 pp.
- Hairston, N.G. Sr. (1989). Ecological experiments. Purpose, design, and execution. Cambridge University Press. Cambridge. 370 pp.
- **Hurlbert, S.H. (1984). Pseudoreplication and the design of ecological field experiments. Ecological Monographs, 54: 187-211.**
- James, F.C. & McCulloch, C.E. (1985). Data analysis and the design of experiments in ornithology. En: Current Ornithology, 2, chapter 1: 1-63.
- Kuehl, R.O. (2001). Diseño de experimentos. Thomson Learning. México. 666 pp.
- Manly, B.F.J. (1992). The design and analysis of research studies. Cambridge University Press. Cambridge. 353 pp.
- Pukelsheim, F. (1993). Optimal design of experiments. John Wiley & Sons, Inc. NY. 454 pp.
- Resetarits, W.J. Jr. & Bernardo, J. (1998). Experimental Ecology. Issues and perspectives. Oxford University Press. NY. 470 pp.
- **Ruxton, G.D. & Colegrave, N. (2003). Experimental design for the life sciences. Oxford University Press. NY. 114 pp.**
- Scheiner, S.M. & Gurevitch, J. (eds) (1993). Design and analysis of ecological experiments. Chapman & Hall. 445 pp.
- Shrader-Frechette, K.S. & Mc Coy, E.D. (1993). Method in Ecology. Strategies for Conservation. Cambridge University Press. Cambridge. 328 pp.
- Sutherland, W. (1996). Ecological Census Techniques. Cambridge University Press. Cambridge. 336 pp.
- Underwood, A.J. (1997). Experiments in ecology. Their logical design and interpretation using analysis of variance. Cambridge University Press. Cambridge. 504 pp.

PROGRAMA DE MÁSTER Y DOCTORADO: “Biodiversidad en Áreas Tropicales y su Conservación”

DISEÑO Y METODOLOGÍAS DE INVESTIGACIÓN EN BIOLOGÍA DE LA CONSERVACIÓN

PRINCIPIOS BÁSICOS DEL DISEÑO EXPERIMENTAL

De forma muy general, las investigaciones pueden clasificarse en cuatro clases:

1. **Documental.** Aprendizaje de nuevos hechos a través del estudio de documentos y registros. *Por ejemplo, la consulta de publicaciones, anuarios estadísticos, etc. (Aves acuáticas-embalses).*
2. **De casos.** Estudio cuidadoso y completo sobre el desarrollo y estado de un fenómeno. *Por ejemplo, mitosis, comportamiento de apareamiento, etc.*
3. **Observacional.** Estudio de un fenómeno sin que este sea modificado por el observador y que se subdivide en:
 - Observacional descriptiva. Cuando el propósito es obtener información relevante y precisa sobre una situación existente. *Por ejemplo, un inventario florístico.*
 - Observacional analítica. El objetivo es establecer el tipo de asociación que pudiera existir entre dos o más características, cuya presencia o ausencia es ajena a la voluntad del investigador; con otras palabras el objetivo es la definición de patrones. *Por ejemplo, la asociación entre la presencia de ciertas especies vegetales y la composición química del suelo* (covariación entre variables que interaccionan). Presenta en muchas ocasiones un elevado grado de generalización.
4. **Experimental.** El objetivo es la comprensión de relaciones funcionales, el establecimiento de asociaciones causa-efecto y la comprensión de mecanismos. Se analizan los efectos de la exposición o privación intencionada de un factor bien definido en parte de los elementos del conjunto en estudio. *Por ejemplo, el efecto de diferentes niveles de competencia sobre la producción primaria.*

Estas diferentes clases de investigaciones no son excluyentes entre sí y, en muchas ocasiones, es difícil definir dónde empieza una y termina la otra. En la investigación documental, de casos y observacional descriptiva, el papel de la estadística es muy restringido, porque el propósito fundamental de la investigación es describir una determinada situación; es decir, sólo se consume la primera etapa del método científico. Por el contrario, en la investigación observacional analítica y en la experimental, la estadística es considerablemente más importante y donde se cumplen las demás etapas del método científico, es decir, formulación y prueba de hipótesis.

Por lo tanto, el **conocimiento**, puede tener su origen en observaciones empíricas, sean o no experimentales, o en estudios científicos, basados en un método demostrativo de relaciones causales, capaz de definir verdades objetivables y reproducibles.

El **profesional** ha de ser científico, es decir, conocedor de verdades teóricas e investigador, por lo que se compromete a describir nuevos hechos o ratificar los ya conocidos, para lo que precisa conocer un método de trabajo que le permita actuar de forma regular, sistemática y ordenada, es decir, el método científico.

La **investigación científica** consiste en la búsqueda permanente de la verdad mediante métodos objetivos, adecuados y precisos, es decir, mediante la aplicación del método

científico. Así, la experimentación es un método científico de investigación que consiste en hacer operaciones y prácticas destinadas a demostrar, comprobar o descubrir fenómenos o principios básicos.

Es interesante destacar que la **experimentación** se puede considerar como un arte y como una ciencia. Como arte, por la habilidad necesaria para idear, planear o aplicar un conjunto de técnicas a fin de eliminar causas extrañas (ej.: *test de lucha para ver si diferentes bocas pertenecen al mismo hormiguero*) y realizar experimentos de campo o de laboratorio (ej.: *perfil de bolas estériles para ver percolación*). Como ciencia, principalmente, por la aplicación del método científico y, también, por la aplicación de un conocimiento científico que permita el desarrollo de nuevas técnicas (ej.: *mejora genética, mayor producción agrícola, etc.*).

Pero ¿qué es el método científico?

Se entiende por método científico o diseño experimental a la estructura lógica de una investigación o de un experimento. El método científico consiste en la búsqueda de hechos, la formulación de hipótesis y la obtención de los principios y las leyes que rigen tales hechos, de manera que se va conformando el cuerpo teórico de una ciencia (ej.: *Hecho = caída de los cuerpos, Hipótesis = son atraídos por la masa de la Tierra, Ley = Gravitación Universal*). Comprende dos aspectos importantes:

- a) **Método inductivo** (de lo particular a lo general). Mediante este método se buscan hechos particulares a través de la observación y la experimentación, de donde se extraen datos a partir de los cuales se intenta llegar a la identificación de patrones y al establecimiento de unos principios generales.
- b) **Método deductivo o hipotético-deductivo** (de lo general a lo particular). Consiste en clasificar y ordenar los hechos por medio de una relación. Si esta relación es constante, de manera que se pueda predecir un hecho y confirmarlo mediante la experimentación, puede ser generalizada y formarse así un postulado básico o una ley. Es decir, se va de un principio universal a hechos particulares, mediante asunciones (generalizaciones) y el estudio de mecanismos (procesos), confirmándose por experimentación y modelos matemáticos.

Cuando se habla del cuerpo teórico de una ciencia se hace referencia a un conjunto de principios reconocidos que pueden aplicarse para predecir una amplia variedad de características específicas de algunos fenómenos generales. Estos principios generales se apoyan en un gran número de pruebas y en la síntesis de los datos y observaciones sobre la organización de la naturaleza. La teoría es una herramienta científica poderosa, pues el razonamiento deductivo a partir de los principios generales permite predecir un gran número de situaciones específicas, muchas de las cuales pueden no haber sido solucionadas antes del planteamiento de la teoría. *Por ejemplo, se puede predecir la fuerza de atracción entre dos cuerpos, la gravitación, según la ley de gravitación de Newton: $F=C(m_1*m_2)/r^2$ (F =Fuerza de atracción, m_1 y m_2 =masa de cada cuerpo, C =Constante de gravitación, r =distancia entre los dos cuerpos).*

La Ecología, sin embargo, no tiene un cuerpo teórico bien conformado como el que caracteriza ciencias como la Física. Los ecólogos todavía dependen en buena parte del razonamiento inductivo, formulando generalizaciones a partir de un número limitado de casos específicos. El desarrollo de una teoría se basa en unos resultados concordantes obtenidos al probar repetidamente una hipótesis específica, y la Ecología, sencillamente, no tiene estas pruebas para un gran número de fenómenos. Cuando se carece de una teoría donde quedan recogidos los términos exactos, matemáticos, el razonamiento deductivo puede ser imperfecto

ya que los científicos a veces usan datos que no prueban en realidad las hipótesis que se habían propuesto demostrar.

En Ecología es interesante tomar leyes o principios de otras disciplinas como por ej. *Las Leyes Termodinámicas*.

Pasos en el método científico y ejemplo

Como en todas las ciencias, la resolución de un problema en ecología es un proceso cíclico:

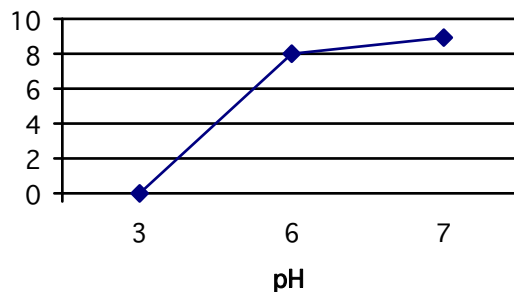
- 1.- Se realiza una observación sobre cierto fenómeno.
- 2.- Por un razonamiento inductivo (de lo particular a lo general) se formula una hipótesis para explicar la observación. Esta hipótesis puede ser un planteamiento verbal o tener un desarrollo matemático fijando ciertas suposiciones sobre el fenómeno.
- 3.- Por un razonamiento deductivo (de lo general a lo particular) se obtiene una predicción a partir de la hipótesis.
- 4.- Se comprueba la predicción, generalmente mediante métodos que dan una descripción cuantitativa de una nueva descripción. Así, la hipótesis inicial puede juzgarse como correcta o incorrecta.
- 5.- Si la hipótesis resulta incorrecta, se formula otra hipótesis basada en la nueva información y el proceso se repite.

Ejemplo 1: Relación entre densidad de una planta y el pH del suelo.

1. Observación: La densidad de una planta depende del pH.

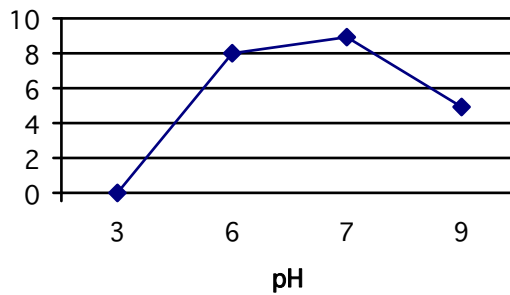
pH = 3 ausencia pH = 6 → 8 individuos pH = 7 → 9 individuos

2. Hipótesis demostrable:
 - Verbal: A mayor pH mayor densidad de individuos
 - Matemática:



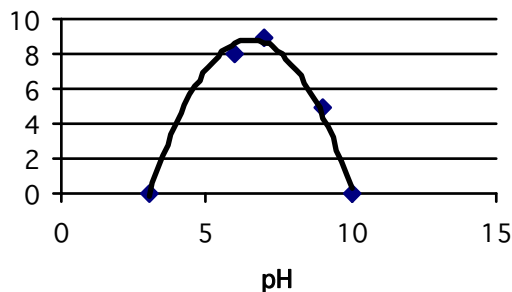
3. Predicción (razonamiento deductivo): Para un pH = 9 la densidad será mayor.
4. Comprobación de la predicción: Experimento con diferentes pH de 3 a 11 por ej.

Resultado pH = 9 → Densidad = 5 individuos



5. Nueva hipótesis:

- Verbal: La planta tiene unos límites de tolerancia
- Matemática:



Ejemplo 2: Con orugas de mariposa Papilio (del McNaughton)

1. Observación: Las larvas se alimentan comúnmente de plantas de la familia de las umbelíferas y raramente sobre otras plantas.
2. Razonamiento inductivo: ¿Pueden estas orugas discriminar entre estas umbelíferas y otras especies de plantas? Los mecanismos de elección podrían ser la forma, el gusto, el olor o el tacto de las hojas.
Hipótesis demostrable: Estas orugas identifican el alimento por el olor que despiden las plantas.
3. Predicción (razonamiento deductivo): Si el olor de las umbelíferas puede transferirse a un material que no es alimento, la larva se sentirá atraída hacia él.
4. Comprobación de la predicción: Triturar varias plantas umbelíferas y empapar con los extractos unos papeles de filtro, colocar tela metálica sobre las muestras y dejar que las orugas caminen sobre dicha tela. Si la visión es importante, las larvas no reconocerán el papel de filtro. Por su parte, la tela metálica elimina las posibles señales de tacto y gusto. Se encontró que las larvas eran atraídas al papel de filtro empapado con los extractos de plantas.
5. Hipótesis correcta: Se demuestra la hipótesis y se dan pruebas, además, de que el mecanismo de reconocimiento es el olor. La experimentación de una hipótesis conduce generalmente a la formulación de otras nuevas. Así, se propone que son

determinados compuestos químicos de las plantas a los que respondían las orugas y que la composición general de la planta era menos importante.

CONCEPTOS GENERALES EN EL DISEÑO EXPERIMENTAL

El término diseño experimental no hace referencia solamente al típico trabajo de laboratorio, es un concepto más general. Se puede definir como la búsqueda planeada para obtener nuevos conocimientos o para confirmar o no resultados de experimentos previos o probar una hipótesis formulada, lo que ayuda a la toma de decisiones.

Etapas fundamentales de la experimentación

Para cualquier experimentación científica son necesarios y en el orden que se presentan los siguientes requerimientos:

- 1) Reconocimiento de que un problema existe.
- 2) Formulación del problema.
- 3) Establecimiento de los factores y niveles que serán utilizados en el experimento.
- 4) Especificación de las variables que serán medidas.
- 5) Definición de la inferencia espacial y temporal para el problema.
- 6) Selección al azar de las unidades experimentales.
- 7) Asignación de los tratamientos a las unidades experimentales.
- 8) Perfilar los análisis correspondientes al diseño antes de que los datos sean tomados.
- 9) Toma los datos.
- 10) Análisis de los datos.
- 11) Conclusiones.
- 12) Implementación.

1) Reconocimiento de que un problema existe. (Planteamiento de la pregunta)

El primer paso a dar en todo diseño experimental es, por encima de todo, tener muy claro cual es su finalidad. Todo trabajo de investigación se debe iniciar concretando el fin perseguido, pero cuidando que esta finalidad no sea demasiado amplia. Previamente al desarrollo del trabajo hay que pensar en los medios materiales y en el tiempo de que disponemos para ajustar los fines perseguidos a los medios disponibles. El principio de la conservación de la energía puede ser aplicado a los estudios estadísticos como norma de conducta. Así, el aumento del número de materias sobre las que pretendemos obtener información nos llevará a perder seguridad en nuestras deducciones. Cada fase de la investigación, así como el tipo de materia de nuestro estudio, nos llevará a decidir si estamos interesados en saber más de pocas cosas, o menos de muchas; los medios de que dispongamos nos marcarán el nivel de ese producto hipotético.

- Fase libre e imaginativa, con el objetivo de plantear la resolución de problemas nuevos, no abordados anteriormente (proceso de investigación).
- Apoyo. En todo trabajo científico siempre hay dos niveles de supuestos: **los que admitimos sin prueba** en nuestro trabajo y **aquellos que vamos a probar en la experiencia**. Cuanto mayor sea el grado de conocimiento previo que tengamos sobre el problema, bien por

experiencias propias anteriores o bien por el estudio de experiencias realizadas por otros investigadores (bibliografía y los caminos de acceso a ella), colaboraciones o consultas a otros investigadores, los supuestos de nuestro trabajo serán más reales, pudiendo realizar estudios más complejos y diseños más sofisticados (es importante saber inglés a nivel de traducción). El número de trabajos científicos que se publican anualmente es cada vez mayor; lo más recomendable es acudir primero a revisiones generales recientes y tras hacerse una idea general del tema que nos interesa, se tratará de ir leyendo artículos y libros más especializados, que a su vez se irán sacando de las referencias bibliográficas que deben acompañar a todo trabajo científico. Muchos investigadores piensan que unos diseños son *per se* mejores que otros, incluso se tiende a pensar que cuanto más complicado es el diseño mucho más perfecto es el estudio y más exactos los resultados. Ambas ideas son totalmente erróneas. Además, no pueden darse reglas generales válidas para siempre; si las hubiera la investigación sería un trabajo mecánico, y no lo es. (proceso de investigación)

- Planteamiento de cuestiones (proceso de investigación).
- Definición de objetivos (proceso de investigación).
- Planteamiento de la forma más clara y explícita posible qué pregunta se quiere responder (proceso de investigación).
- Identificación del problema o problemas pero, y esto es muy importante, que sean **analíticamente abordables** (apoyo de la estadística).
- Formulación de hipótesis del método hipotético deductivo (apoyo de la estadística).
- Enunciado de predicciones (apoyo de la estadística).

Esta fase de reconocimiento de un problema en ecología puede no ser muy larga, más aún si se tiene en cuenta que es una ciencia relativamente reciente, pero es muy importante que quede perfectamente clara y definida.

Planteamiento de la pregunta

Preguntas sencillas, concisas y directas

Hay preguntas, que por lo general se pueden contestar con más o menos dificultad, como es el caso de: ¿Cuál? ¿Cuánto? ¿Dónde? ¿Cuándo? ó ¿Cómo?

Pero, en cambio la pregunta ¿Por qué? no es la más apropiada para iniciar una investigación. Veamos un ejemplo.

Ejemplo. Tenemos un bosque con la categoría de reserva con dos tipos de fronteras:

1. Bordes del bosque que lindan con vegetación secundaria (matorral).
2. Bordes del bosque que lindan con pastos.

Además existe una planta invasora que está ocupando la reserva.

Ante esta situación yo me puedo plantear la siguiente pregunta:

¿Cómo difiere el número de plántulas de la especie invasora en cada borde?

Esta pregunta puede ser fácilmente abordable. Pero si la pregunta fuera:

¿Por qué hay más plántulas en el primer borde que en el segundo?

Contestas a esto implica que, además de tener que haber más plántulas en el 1º borde que en el 2º, en el momento presente esta pregunta no puede contestarse directamente pues este resultado involucra eventos del pasado que pueden influir en la llegada, supervivencia y germinación de las semillas de la especie invasora como son: el efecto de las aves y mamíferos frugívoros, los usos del suelo, los ataques de hongos o patógenos a semillas, la prelación por ratones y hormigas, las propiedades del suelo, el microclima, la exposición al viento y, como no, el azar.

Por lo tanto, la pregunta ¿Por qué? es preferible considerarla en la fase de reflexión (interpretación de los resultados).

Preguntas pomposas = preguntas absurdas

Por ejemplo: ¿Las agregaciones transitorias de *Gallus gallus* cimarrones que están asociadas a establecimientos agrícolas de subsistencia adyacentes, afectan negativamente la supervivencia de los propágulos y el reclutamiento de los juveniles de la vegetación arbustiva nativa, dentro de un módulo de manejo, comparado con parcelas control de exclusión?

Pregunta más sencilla (útil)

¿Cómo varía la pérdida de semillas y plántulas de especies nativas, y la regeneración del bosque, entre áreas con presencia y ausencia de gallinas?

Preguntas comparativas

Este tipo de preguntas son de enorme utilidad para llevar a buen fin una investigación, puesto que requieren que el investigador tenga que enfocar el problema (ponerlo en contexto) y, además le permite una reflexión posterior, como ahora veremos. En cambio, una pregunta no comparativa es casi siempre un callejón sin salida y no permite una reflexión posterior ni da opción de seguir avanzando en la investigación. Vamos a analizar estas afirmaciones con algunos ejemplos.

Pregunta 1. ¿Cuántas especies de primates viven es una mancha de bosque, y cuáles son raras y están amenazadas?

Esta es una pregunta no comparativa, por supuesto que es válida pero ¿qué utilidad tiene aparte de describir el estado de un fragmento del bosque?

Pregunta 2. ¿Hay diferencias en el número de especies de primates raras y amenazadas entre dos manchas de bosque del mismo tamaño?

Esta pregunta si es comparativa pero no llega mucho más lejos que la anterior, pero eso sí con el doble de trabajo. Pero también, como la anterior, le falta un contexto, un factor general que permita al investigador dar las bases para tomar decisiones de manejo respecto al resto de los fragmentos dispersos por la reserva.

Pregunta 3. ¿Cómo varía el número de especies de primates raras y amenazadas entre una mancha de bosque pequeña y otra grande?

En este caso, la pregunta es comparativa y tiene un contexto (efecto del tamaño de la mancha en el numero de especies), además incorpora una clave para posibles decisiones de manejo. Pero hay un problema, y es que si hay diferencias entre las manchas puede que no se deban a su diferente tamaño sino a características particulares de los fragmentos (diferente densidad de árboles o matorral, diferente meso o microclima...).

Pregunta 4. ¿Cómo varía el número de especies de primates raras y amenazadas entre fragmentos de bosque de diferente tamaño?

Esta pregunta es la más sugerente y la más útil para la toma de decisiones de manejo, pues aquí esta considerada toda la reserva. Pero incluso esta pregunta podría decirse que es una pregunta tonta, puesto que todos sabemos que una tendencia general es que los fragmentos más grandes llevan asociados un mayor número de especies. Pero en general, en nuestro caso, hay que demostrarlo. Podría ocurrir que los fragmentos pequeños presenten un entorno caracterizado por un matorral secundario en regeneración que brinde muchos más recursos y permita un mayor número de especies de primates que los grandes.

Para seguir con el experimento es necesario plantearse qué tipo de información necesitamos para contestar nuestra pregunta, es decir, que cosas vamos a medir y dónde y cuándo vamos a medirlas.

2) Formulación del problema.

Una regla (que evidentemente es tan ambigua como problemas o investigaciones diferentes existen) es que toda investigación debe responder a cinco preguntas:

1. *Qué (Hipótesis)*. Finalidad clara y precisamente formulada.
2. *Quién (Objeto de estudio)*. Individuo, población, comunidad, factores, variables, etc.
3. *Cómo*. Plan de trabajo, aparatos de medida, escalas, toma de muestras, coste, etc.
4. *Dónde*. Zona, laboratorio, clima, efectos locales, homogeneidad del terreno, etc.
5. *Cuándo*. Estacional, anual; según edad, densidad, perturbación, etc.

Cada una de las contestaciones a estas preguntas debe concretarse en función de la primera: la finalidad del trabajo.

A modo de resumen, se podría decir que la abundancia de medios, la mayor cantidad de datos tomados como sea, los más sofisticados diseños, los más potentes medios informáticos, no implican un buen trabajo de investigación.

El problema en cuestión puede estar originado por un elevado número de causas (factores). Es muy importante reducir el número de factores, previa selección de los más relevantes.

3) Establecimiento de los factores y niveles que serán utilizados en el experimento.

Factores

Se llama factor a aquella variable experimental que es independiente, es decir, que el investigador controla perfectamente y de la que se quiere estudiar su efecto, sobre ciertas características de la unidad experimental o de muestreo. Los factores son fijados por el investigador, bien en el campo (*ej.: pendiente, grado de insolación, densidad de individuos, etc.*) o en laboratorio (*ej.: concentración de nutrientes, temperatura, superficie de forrajeo, etc.*). Los factores pueden ser **cuantitativos**, cuando toman valores correspondientes a los puntos de una recta real (*ej.: peso, temperatura, longitud recorrida, etc.*) o **cualitativos** (*ej. color de la flor, tipos de comportamiento, sexo, etc.*).

Cuando dentro de un factor se consideran diferentes grados de intensidad, a cada uno de estos grados se les denomina **niveles del factor**. *Por ejemplo, Factor: sexo, niveles: macho y hembra; factor: [nitrito], niveles: 1 g/l y 0,1 g/l.*

Tratamiento

Generalmente se confunde con factor, puesto que en muchas ocasiones es lo mismo. Pero hay que tener en cuenta que el concepto de tratamiento es más amplio, pudiendo definirse como la combinación específica de niveles de un factor o de factores. Por ejemplo, en la práctica de producción primaria un tratamiento es: alta competencia (50 semillas)-baja [nitrate] (agua).

Una vez seleccionados los factores más relevantes, hay que encontrar los niveles reales en los que opera este factor que nos permita detectar el efecto de los diferentes factores. Por supuesto, la determinación de los niveles óptimos se decidirá al final de la investigación experimental. Ej. En el ejemplo visto antes del efecto del pH sobre la densidad de una especie vegetal, sería absurdo estudiar todos los niveles de pH (1, 2, 3....14). O bien en la práctica de los alumnos con los niveles de los factores: $[NO_3]$ y competencia.

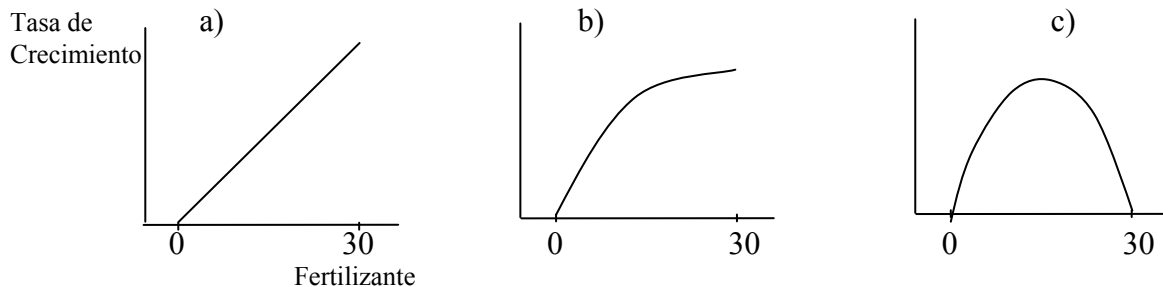
Ejemplo: Ver el efecto de un fertilizante sobre la tasa de crecimiento de las plantas de tomate ($n=40$ plantas).

1. Supongamos que hacemos 2 niveles del factor

- Nivel 1: 0 mg
- Nivel 2: 30 mg

Dos grupos de tratamientos extremos cada uno con 20 plantas (máximo poder de análisis)

2. Supongamos que hay 3 posibles relaciones entre cantidad de fertilizante añadido y el incremento en la tasa de crecimiento:



Caso a). El resultado tiene gran poder estadístico y refleja la realidad. La pendiente que se deduce es correcta.

Caso b). Ofrece información incompleta. No se detectaría la importancia biológica (y probablemente económica) que implica que un pequeño incremento de fertilizante a bajos niveles tiene una gran efecto sobre la tasa de crecimiento, mientras que a altos niveles un pequeño incremento tiene un pequeño efecto en la tasa.

Caso c). Se concluye que el fertilizante no tiene efecto, cuando es todo lo contrario.

Solución: Si no se conoce la relación (lo normal) es necesario tener grupos de tratamientos (niveles) intermedios. Esto reduce el poder estadístico (4 niveles = 10 individuos por nivel) pero ofrece resultados más reales.

Los datos de los estudios cuyos factores de diseño tienen niveles categóricos (discretos, continuos, cualitativos), a menudo se analizan por medio del análisis de la varianza (ANOVA). Este tipo de factores tiene la ventaja de que permite investigar más de un factor al mismo tiempo y los estudios son relativamente fáciles de realizar e interpretar y el poder detectar efectos estadísticos es muy grande. La desventaja es que si se hace una investigación

experimental y se eligen ciertos niveles podría suceder que en la realidad el factor ejerza sus efectos más destacados en los niveles que no se han elegido, bien más extremos o intermedios.

Los datos de los estudios cuyos factores de diseño tienen niveles continuos (cuantitativos), a menudo se analizan por medio de análisis de regresión.

4) Especificación de las variables que serán medidas.

Las variables son aquellos parámetros o aquellas características que el investigador selecciona de la unidad experimental o unidad de muestreo y cuya medida informa de los posibles efectos que sobre ellos (variables dependientes) originan las variables independientes o factores (causa). *Por ejemplo, peso de la raíz, diversidad de especies, período de actividad, etc.*

Una variable puede definirse como toda aquella característica que difiere entre unidades de muestreo. En esta fase se analizan las variables relevantes para el examen de hipótesis y su independencia o no, es decir, las respuestas de las variables a un factor pueden ser independientes las unas de las otras y, consecuentemente, pueden ser tratadas como variables dependientes separadas. *Por ejemplo, para estudiar la tasa de crecimiento hay que elegir muy bien la variable, ya que puede ocurrir que la elegida (grosor del tallo) no cambie con la variación del factor.* Por el contrario, las respuestas de las variables frente a un factor puede ser dependientes y habrá que tratarlas conjuntamente (covariables).

El valor de una variable varía entre observaciones. Esta variación es extraordinariamente importante en estadística y es muy importante medirla adecuadamente. Es muy importante usar un método de medida adecuado para recoger esta variabilidad. *Así, por ejemplo, si tomamos los siguientes datos en un encinar:*

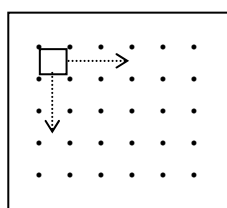
Var1(altura de las encinas): 13, 13.5, 13.2, 13.7, 13.1 (metros)

Var2(nº de encinas): 20, 25, 30, 19, 37

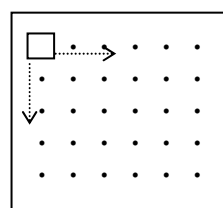
Var3(presencia-ausencia): sí, no, no, no, sí

Sí por ejemplo, hubiésemos medido la altura de las encinas con un método cuyo error de medida es de 1 m encontraríamos que todas miden lo mismo (13 m), a no ser que se siguiesen midiendo árboles hasta encontrar uno superior a 13 metros. Algo similar ocurrirá para la Var2 si el número de encinas contadas fuese igual en todas las superficies de 100 m² medias (como es el caso de una plantación con distribución regular de los individuos), o bien en la Var3 si aparecen encinas en todas las unidades de superficie de 100 m² o no aparecen en ninguna. En estos tres casos, ni la altura ni el nº de individuos ni la presencia-ausencia serían variables en este estudio.

Ej.: Dos territorios con diferentes poblaciones de ratones con clara territorialidad en los que se muestrea de forma regular para ver si hay diferencias en densidad de individuos.



TERRENO 1



TERRENO 2

En ambos tratamientos las unidades de muestreo utilizadas tienen el mismo tamaño y el muestreo aplicado es el regular, comenzando por tomar el punto del primer dato al azar. En el primer terreno, se obtendría que no existe ninguna población de ratones y en el segundo terreno sí.

En función del número de valores que una variable puede tomar, se definen dos tipos:
- **Variables continuas.** Si pueden tomar, al menos teóricamente, cualquier valor entre dos valores dados. Por ejemplo, longitudes, pesos, cobertura, etc. La finura de la medida sólo dependerá del aparato, aunque pueden considerarse también continuos los conteos siempre y cuando el número de resultados posibles sea elevado (en general, más de 20).

- **Variables discontinuas o discretas.** Sólo pueden tomar unos pocos valores en un intervalo dado. Por ejemplo, conteos con pocos resultados posibles (menos de 20, Ej.: Variable: *Número de puntas en la cornamenta de los ciervos.*) y las categorías (sexo, edad, grande/pequeño, arriba/abajo, presencia/ausencia, etc.), por lo que también se denominan **variables categóricas**.

Es muy importante distinguir los tipos de variables con que se está trabajando de cara a su análisis estadístico, puesto que las pruebas que se usan para cada tipo son diferentes.

Dificultad en la elección de las variables

En muchas ocasiones hay pocas dudas de la variable que se debe medir. Sin embargo, a menudo la elección no está tan clara.

Ejemplo: *Supongamos que se estudia la invasión del bosque por maleza después de una tala selectiva, y se han localizado áreas con invasión de maleza en zonas sin tala (ST) y en zonas con tala selectiva (TS).*

La pregunta sería: ¿Cómo varía la tasa de invasión de malezas, entre parcelas TS y ST?

Variable a medir: número de plántulas de especies invasoras. Pero no sería mejor estas otras:

- *Proporción de plántulas de especies invasoras con respecto al total de plántulas. Podría ocurrir que en las áreas de TS hay mayor entrada de luz y puede aumentar la germinación tanto de las especies nativas como de las invasoras.*
- *Tasa de crecimiento de un determinado número de plántulas durante un cierto tiempo.*
- *Seguimiento de las tasas de supervivencia de plántulas, considerando que la mayoría morirá antes de alcanzar un tamaño que tenga consecuencias para la conservación.*
- *Número de especies de plantas invasoras.*

Cada elección de una de estas variables nos lleva a una respuesta cualitativamente diferente y por lo tanto a distintas reflexiones, conclusiones y decisiones de manejo.

5) Definición de la inferencia espacial y temporal para el problema.

Universo de muestreo. Todos los posibles objetos sobre los que se pueden tomar mediciones. Grupo al que se quiere inferir.

Unidad experimental. Los objetos sobre los cuales se hacen mediciones se denominan unidades experimentales. O bien, las unidades a las que se les aplican los tratamientos (factores). *Por ejemplo, un individuo, una población, una ladera, etc.*

Unidad de muestreo. Es la fracción de la unidad experimental a la que se le hacen mediciones.

Ejemplos:

- 1) *¿ [melanina] en los españoles?. Se usa un aparato que mide la [melanina] en 1 cm² de piel.*

Universo muestral: Todos los españoles.

Unidad experimental: el individuo.

Unidad de muestreo: 1 cm² de piel.

- 2) *En el mismo ejemplo, yo considero que el universo muestral puede ser muy heterogéneo para la variable a estudiar. Por ejemplo, podría encontrar grandes diferencias en la variable entre N y S de España, entre Comunidades Autónomas, entre zona Rural y Urbana, o entre Hombres y Mujeres. En este caso:*

Universo muestral: Todos los españoles.

Unidad experimental: Zona de España, Comunidad Autónoma, Rural o Urbano, sexo y el individuo.

Unidad de muestreo: 1 cm² de piel.

Puede ocurrir que la unidad de muestreo coincida con la unidad experimental.

- 3) *¿Cuál es el peso medio de los españoles?*

Universo muestral: Todos los españoles.

Unidad experimental: el individuo.

Unidad de muestreo: el individuo.

En este caso coincide la unidad experimental con la unidad de muestreo.

- 4) *¿Cuál es la altura media de los nidos de urraca en un encinar (cerca y lejos de la carretera).*

Universo muestral: Todos los árboles del encinar con nido.

Unidad experimental: el árbol con nido.

Unidad de muestreo: el árbol con nido.

Podría ocurrir que por comodidad todos los árboles que se midan estén cerca de las carreteras, lo que provocaría un sesgo en los datos (cerca de la carretera los nidos están situados a mayor altura que en los árboles que están en el interior del bosque).

Pero, por lo general, en Ecología la unidad experimental es lo suficientemente grande y compleja que su uso como unidad de muestreo no es práctico. En este caso se toman varias unidades de muestreo de pequeño tamaño, constituyendo lo que se denomina **muestra**. *Por ejemplo, en el estudio de la dinámica de la comunidad de invertebrados epiedáficos en una ladera entre la parte alta y la baja, es conveniente situar varias trampas de caída arriba y otras tantas abajo, constituyendo cada grupo **una muestra**.*

Una vez que se ha definido el problema, se pueden tomar decisiones sobre los límites de las inferencias que se hacen a partir de los resultados. Así, los límites dentro de los cuales los resultados serán aplicados es lo que se podría denominar inferencia (espacial o temporal). *Por ejemplo, si un químico quiere que los resultados de su experimento puedan aplicarse o extenderse a todos los laboratorios del mundo, es necesario que sean incluidos en el experimento un número lo suficientemente grande de muestras tomadas al azar de los posibles laboratorios.*

6) Selección al azar de las unidades experimentales.

Por definición, **en la naturaleza nunca existen dos elementos iguales**. El término unidad experimental hace referencia al tipo de material experimental que va a recibir la aplicación de los diversos tratamientos (factores) y que será representativa de la inferencia espacial deseada. *Por ejemplo, si uno está interesado en investigar los efectos de varios tipos de fertilizantes sobre plantas en tiestos, la unidad experimental será la planta en el tiesto*: en este caso hay que tener mucho cuidado antes de extender la inferencia a condiciones de campo. *Por ejemplo, en ecología se utiliza muchas veces un cuadrado de suelo como unidad experimental*: en este caso, el tamaño del cuadrado es normalmente dependiente de los hábitos de crecimiento de la planta en cuestión porque los resultados deben ser aplicados a la población deseada. *Ej: Datos censurados en Etología.*

Una vez que ha sido definida la unidad experimental el investigador tiene que seleccionar al azar un número suficiente de unidades para ser utilizadas en el experimento. La selección al azar es necesaria para evitar cualquier sesgo en nuestro experimento, ya que de lo contrario los resultados podrían estar reflejando el efecto de algún factor desconocido teniendo una influencia previa sobre las unidades experimentales de alguna forma sistemática.

Si se emplean criterios de aleatorización y en consecuencia se aplicarán tests de significación (estadística inferencial), la replicación es obligatoria al nivel de unidad experimental.

El número de unidades experimentales es claramente dependiente del error estándar deseado sobre la media. Así, en ciertas investigaciones las unidades experimentales disponibles no son homogéneas lo que obliga a seleccionar al azar un mayor número de éstas.

Esta bien demostrado que cuando la variable a medir está expresada por una unidad de espacio, de tiempo, de esfuerzo o por un número determinado de individuos, diferentes tamaños de unidad produce resultados y conclusiones completamente diferentes. Ahora bien, hay pocos recursos para solucionar este problema, sólo queda el sentido común y el conocimiento de la historia natural.

Ejemplo: Pocas trampas de caída para ratones en el espacio y el tiempo. Se obtendrían muchos ceros. No refleja la realidad, se miden eventos netamente aleatorios.

7) Asignación de los tratamientos a las unidades experimentales.

Esta fase es lo que se denomina diseño experimental. Muchas veces debido a la naturaleza del experimento, las restricciones sobre la azarosidad de la aplicación de los tratamientos a las unidades experimentales están inherentes (a veces muy hábilmente disfrazadas) en el diseño, es decir, el investigador piensa que los tratamientos están siendo adjudicados al azar y realmente no es así.

La asignación aleatoria de los tratamientos a las unidades experimentales simula el efecto de independencia y permite proceder como si las observaciones fueran independientes y con distribución normal. Esta aleatorización constituye la base de la inferencia estadística.

Por ejemplo, quiero ver el efecto de un fertilizante sobre el crecimiento (variable) de una especie de planta respecto a dos sustratos (1,2), para lo cual selecciono al azar los individuos a los que les aplico el tratamiento; si primero hago el experimento para un sustrato y luego el del otro, ocurrirá que si hay diferencias en el crecimiento, puede ser debido efectivamente a los diferentes sustratos o a otro factor que no se controla como el hecho de que las condiciones ambientales presentes durante el experimento del primer sustrato sean diferentes a las del segundo.

Ej: Dos diferentes formas de abonar

NO																			
A1	A1	A1	A1	A1	A1	A1	A1	A1	A1	A1	A2	A2	A2	A2	A2	A2	A2	A2	
SI																			
A1	A2	A1	A1	A2	A2	A2	A1	A2	A2	A1	A2	A1	A1	A2	A2	A1	A2	A2	A1

También se podrían aplicar alternados, siempre que no coincida con una regularidad en el medio.

8) Perfilar los análisis correspondientes al diseño antes de que los datos sean tomados.

Antes de seguir con el proceso experimental hay que ver si existen tests estadísticos que se puedan aplicar a nuestro experimento y que las suposiciones (en orden a testar los efectos de una o varias variables) y conclusiones son realistas. En esta fase es importante investigar diversos diseños, llegándose a un compromiso entre minimizar los costes (económicos, de tiempo, de esfuerzo, etc.) de llevar a cabo un determinado experimento, las restricciones sobre la randomización y los posibles gastos como consecuencia de tomar una decisión incorrecta en cuanto que el experimento elegido no permite realizar un test que explique el efecto del factor a estudiar sobre la variable medida.

9) Toma los datos.

Una vez que se tiene claro qué vamos a medir, dónde y cuándo medirlo, hay que tratar de aplicar el método de medida correcto. Hay que establecer el tipo de muestreo más idóneo para la toma de datos. Existen muchos manuales publicados sobre este aspecto. También es importante saber cuándo parar de tomar datos.

Para los investigadores inexpertos la colección de datos es un tópico particular con un procedimiento simple y claro. Por el contrario, hay que tener mucho cuidado. Muchas veces llevar a cabo un experimento puede requerir un equipo organizado con muchos observadores. El conjunto de sucesos de **una investigación científica depende de la validez de todos los datos obtenidos**. El tiempo de muestreo incrementa la experiencia del muestreador (sesga la información).

Ej. Estimación de la diversidad de anfibios en 20 lagunas. Resultados: Las más diversas son las 5 últimas muestreadas, cuando es falso (se mide el aprendizaje del muestreador y no la diversidad).

Ej. Los agujeros de las cápsulas de las jaras se consideran de salida cuando realmente son de entrada; esta acepción implicaría que se considera que el adulto selecciona el fruto, cuando realmente es la larva.

10) Análisis de los datos.

La interpretación de cualquier fenómeno biológico requiere tener en cuenta tres aspectos fundamentales:

- 1.- **Carácter variable de las manifestaciones.** Uno de los rasgos más extendidos de los fenómenos biológicos es que sus manifestaciones no se repiten exactamente, de manera que las mediciones que se obtienen nunca son completamente iguales.
- 2.- **Información incompleta.** Una situación muy generalizada es la imposibilidad de observar todas las posibles manifestaciones del fenómeno estudiado. *Ej.: Mecanicismo.*
- 3.- **Causas múltiples.** Los fenómenos biológicos, además de estar afectados por las causas analizadas en nuestro estudio, están además afectados por otros muchos factores que es necesario considerar a la hora de interpretar los resultados obtenidos. De hecho, *esta es la principal razón de la variabilidad que encontramos en el mundo natural.*

La herramienta que nos permite manejar e interpretar los datos y que a su vez tiene en cuenta estos problemas es la estadística. Esta disciplina tiene dos acepciones:

- Conjunto de métodos empleados para la exposición ordenada y resumida de datos cuantitativos.
- Conjunto de métodos que permiten establecer inferencias sobre las propiedades de grupos grandes a partir de observaciones en grupos pequeños.

La estadística, de esta forma, permite resumir los datos obtenidos teniendo en cuenta su variabilidad y los objetivos de nuestros estudios, así como establecer resultados generales a partir del estudio de conjuntos de observaciones particulares. Tener un dominio de los conceptos estadísticos básicos y de unos cuantos métodos de análisis nos permitirá diseñar correctamente nuestro estudio evitando el efecto de factores indeseados, establecer el grado de fiabilidad de nuestros resultados de modo que puedan ser publicados, y leer y evaluar crítica e inteligentemente los resultados publicados por otros investigadores.

Esta fase es por lo tanto principalmente responsabilidad del estadístico. Evidentemente el análisis de los datos es muy dependiente del diseño del experimento. La representación gráfica es una buena vía para expresar los resultados.

En resumidas cuentas, en esta fase se realiza un **contraste de los resultados obtenidos en relación con lo que cabría esperar por azar.**

11) Conclusiones.

En la mayoría de los casos el investigador será capaz de interpretar físicamente los resultados estadísticos que han sido puestos a la luz por el experimento. Es importante tener en cuenta la magnitud de los resultados y los costes de la implementación de algún nuevo procedimiento que pueda ser requerido.

Una vez interpretados los resultados se lleva a cabo una propuesta de conclusiones y generalizaciones. Esta fase nos lleva a la identificación de posibles interacciones desconocidas y a la reformulación o formulación de nuevas hipótesis que podrán ser el punto de origen de nuevas investigaciones.

12) Implementación.

En esta fase se demuestra que con la investigación científica realizada se ha desarrollado un tipo realista de experimento, se asegura el propio desarrollo de éste y se encuentran resultados seguros.

ANÁLISIS DETALLADO DEL DISEÑO EXPERIMENTAL

Representatividad, Independencia y Factores Alineados (Muestreo)

La selección de las unidades de muestreo que componen nuestra muestra se denomina genéricamente **muestreo**. Nuestra muestra debe ser **representativa del universo muestral** para poder extrapolar los resultados que obtengamos de su análisis. Entre los principales problemas que competen a esta representatividad está la no independencia de las muestras y el efecto de factores no controlados que puedan estar afectando a nuestros resultados. *Por ejemplo, si estamos estudiando los individuos de una población, si las unidades de muestreo son individuos que pertenecen a la misma familia (ej. variabilidad genética), es probable que se parezcan más entre sí con respecto al resto de los individuos de la población de modo que no son representativos del conjunto de la población. Por ejemplo, si medimos la abundancia de una especie en un lugar en invierno y de otra en otro lugar en verano, las diferencias probablemente serán debidas tanto a efectos espaciales como temporales. aspectos que habrá que tener en cuenta al interpretar los resultados.*

La **solución** a estos problemas consiste en una adecuada **replicación** de las muestras, de manera que puedan considerarse homogéneas desde el punto de vista de los factores que las afectan.

Hay que asegurarse de tener verdaderas repeticiones de las unidades de muestreo dentro de cada nivel del factor (unidades experimentales)(factores categóricos). Este paso, y sólo este paso, nos permite distinguir la variación asociada con el factor de la variación intrínseca entre diferentes unidades de muestreo. Por lo tanto, si los niveles son categóricos, las múltiples unidades de muestreo del nivel dado se denominan réplicas. Si los niveles son continuos, por supuesto habrá pocas pocas o ninguna réplica, ya que cada una de las unidades de muestreo puede corresponder a una unidad experimental (factor) única. Obviamente el juego de unidades de muestreo asociado a la gama de unidades experimentales debe ser independiente.

Una replicación adecuada persigue **hacer máximas** dos características importantes del muestreo:

- 1) Su **exactitud (sesgo → representatividad)** o grado de acercamiento al valor real.
- 2) Su **precisión (dispersión)**, que es mayor cuanto mayor es el porcentaje de observaciones coincidentes.

El **método** para hacer máximas estas características es necesario considerar dos aspectos independientes aunque interrelacionados:

- **Qué unidades** deben elegirse de entre todas las posibles.
- **Cuántas** son suficientes para obtener resultados representativos.

Las únicas metodologías de muestreo libres de sesgos, son las usadas para plantas y animales sedentarios. Casi toda la metodología usada para muestrear animales móviles, organismos crípticos, etc., está sesgada en mayor o menor grado, desde el momento en que la población

muestreada difiera sustancialmente de la población real. Lo importante es tener la seguridad que el error que se comete por la metodología de muestreo aplicada sea la misma en todas las unidades experimentales. El problema más grave surge cuando el error que se comete es diferente en las unidades experimentales.

Por ejemplo: Puede ocurrir que la capturabilidad de ratones con trampas de caída en dos áreas diferentes (por ej. Sin Tala y Tala Selectiva) sea muy similar. Pero, en cambio, la capturabilidad de aves con redes abiertas en el sotobosque de ambas áreas puede ser muy diferentes; las redes muestrearían un mayor porcentaje de la avifauna total del bosque en los sitios con TS que en los ST, porque en los de TS hay muchas más zonas aclaradas y un mayor número de aves puede volar cerca del suelo y ser capturadas.

Factores potencialmente alineados

Son aquellos factores que no han sido tenidos en cuenta en la investigación y que pueden exagerar, contradecir, oscurecer o, en cualquier caso, complicar el efecto del factor o factores que estamos analizando sobre los resultados.

Soluciones:

1. Esparcir aleatoriamente las unidades de muestreo por los diferentes niveles del factor en análisis. Esto equivale a minimizar los efectos de los factores realmente alineados sobre los datos (por la fuerza bruta).
2. Controlar los factores más obvios restringiendo el hábito de muestreo (por ej: usando el muestreo estratificado).
3. Implementar un diseño de bloques para los factores potencialmente alineados más obvios. Este tipo de diseño tiene la ventaja de que el análisis estadístico permite separar la variación entre unidades de muestreo que se debe al factor analizado, de aquella debida a los factores potencialmente alineados.
4. Realizar un estudio experimental. (*Manipulación de los factores*)
5. Aceptar la existencia de los factores alineados y replantear la pregunta de manera que se evite esta situación.

Selección de las unidades de muestreo

Si nuestro universo muestral de unidades de muestreo puede considerarse homogéneo, el mejor modo de seleccionar las unidades que vamos a medir es **al azar** de entre todas las posibles, es lo que se denomina **muestreo aleatorio simple**. Para ello se numeran todas las unidades y se extrae un conjunto de números al azar a partir de una calculadora, un ordenador o una tabla estadística. A continuación se localizan las unidades y se miden. Como alternativa puede también realizarse un **muestreo sistemático**, en que se comienza en un punto al azar y luego se van seleccionando las unidades a intervalos fijos de espacio y/o tiempo. Aunque estas unidades no son estrictamente independientes, ya que están a una distancia espacial y/o temporal fija de la anterior y posterior, pueden considerarse como tales si el intervalo no coincide con alguna regularidad del ambiente que pueda afectar los resultados (por ejemplo, animales con territorialidad). *Por ejemplo, si se midiera la producción primaria en árboles caducifolios de invierno a invierno, se perdería información.*

Si nuestro universo de muestreo es claramente heterogéneo, es decir, hay grupos de unidades de muestreo con características diferentes que puedan afectar a nuestros resultados, el mejor modo de seleccionar unidades consiste en dividir primero el universo muestral en grupos homogéneos de unidades o **estratos de muestreo**, en los que luego se realiza un

muestreo aleatorio simple o sistemático. Este procedimiento aumenta sobre todo la precisión del muestreo, ya que agrupa muestras que serán posiblemente más parecidas entre sí.

Estos tres métodos (muestreo aleatorio simple, muestreo sistemático y muestreo estratificado) son los más sencillos y adecuados para replicar correctamente nuestras muestras, ya que las distribuyen de modo homogéneo entre el universo muestral aumentando la exactitud de nuestros resultados. Esto es muy importante ante el posible efecto de factores desconocidos que puedan afectar a nuestro estudio, como es el caso de un factor que varíe de manera sistemática a lo largo del universo muestral. Es evidente, que aunque se realice una adecuada replicación, la precisión de nuestras observaciones puede verse afectada por este factor (aumentará la variación entre nuestras), pero sin duda su exactitud se verá mantenida, pues se sigue recogiendo información de toda el universo muestral (se mantiene la representatividad de la muestra).

En definitiva la **aleatorización** permite la obtención, **sin sesgos**, de un conjunto de datos representativos de las unidades muestrales bajo estudio. Por el contrario, la **no aleatorización** produce **sesgos espaciales o temporales**.

Cuando no conviene aplicar el azar (individuos, espacio, tiempo)

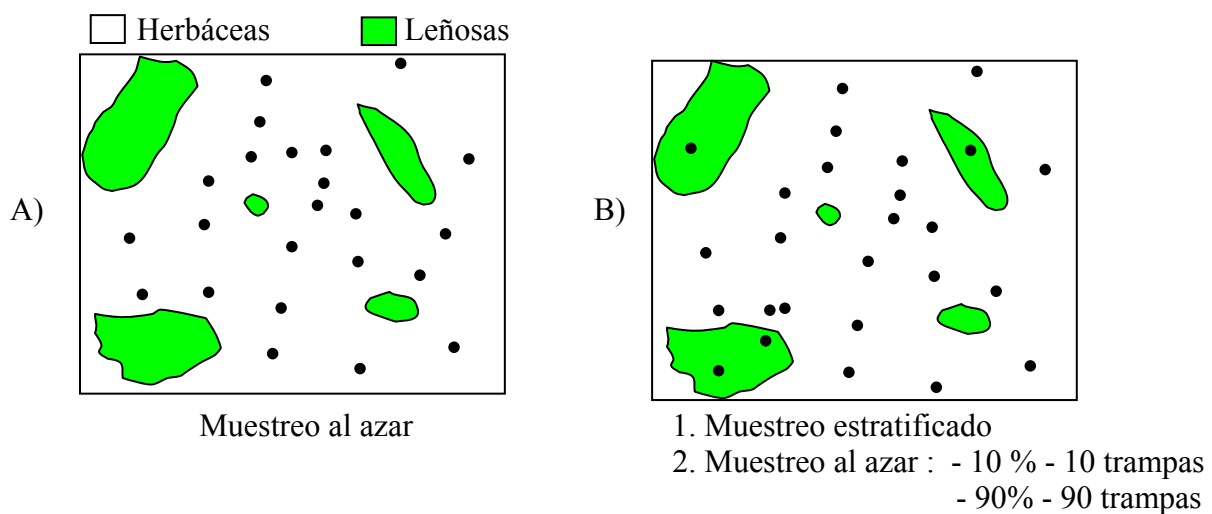
Ej (en el espacio): En el caso de los tomates imaginar que para hacer el experimento se necesitan dos invernaderos y que éstos tienen un efecto sobre la tasa de crecimiento. Lo que no tiene sentido es aplicar un muestreo al azar completo, es decir, adjudicar los individuos al azar en cada invernadero y aplicar al azar a un invernadero 0 mg de fertilizante y al otro 30 mg.

El muestreo se hace por bloques de la siguiente manera:

<u>Invernadero 1</u>	<u>Invernadero 2</u>
10 indiv. → 0 mg	10 indiv. → 0 mg
10 indiv. → 30 mg	10 indiv. → 30 mg

Comparación entre muestreo al azar y estratificado

Ej: Un áreas con dos zonas: 90% ocupado por herbáceas y el 10% ocupado por leñosas. Se colocan 100 trampas para ver la diversidad de hormigas en este bosque fragmentado.



En el caso A) al azar, el área de leñosa no es muestreada, pero nosotros queremos tener información de esta zona (no nos vale el azar)

En el caso B) se hace primero un muestreo estratificado y dentro de cada estrato se muestrea al azar manteniendo la proporción, es decir, sin incrementar el esfuerzo de muestreo.

Pseudorreplicación

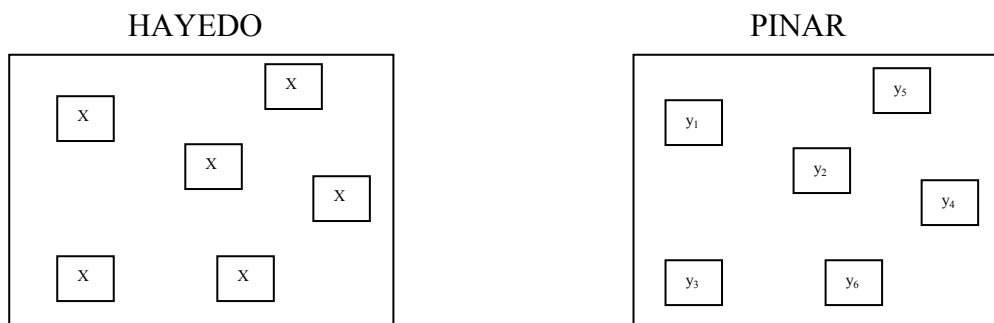
Es muy común que en los trabajos de investigación el modo en que se realizan las observaciones de campo no se llevan a cabo con unas réplicas rigurosas, ya que no hay modo después de asegurar que los resultados que obtenemos se deban a la causa que nosotros estudiamos y no a otra cualquiera (por ejemplo, el factor desconocido que varía sistemáticamente sin que nos demos cuenta). Estas observaciones **no** cumplen la necesidad de **independencia** y se dice que están pseudorreplicadas o que se incurre en **pseudorreplicación**.

Si los tratamientos están segregados espacial o temporalmente, si todas las réplicas de un tratamiento están interconectadas de alguna forma (por ejemplo, por comartición de características ambientales) o si las réplicas son sólo muestras de una única unidad experimental, las réplicas no son independientes. Si se usan los datos de tales experimentos para testar los efectos de los tratamientos se está cometiendo pseudorreplicación → resultados sesgados → menor exactitud → no representativo.

Se dará o no pseudorreplicación dependiendo de la hipótesis.

Pseudorreplicación simple. Es el caso más sencillo. usando una sola réplica. Cuando se hacen medidas múltiples sobre el mismo individuo y se consideran independientes. Se suelen emplear tests de significación anteriores al tratamiento para comprobar que ambas unidades experimentales son iguales. Pero, sin embargo, por definición dos zonas cualesquiera son siempre diferentes, y la falta de significación únicamente se debe a que el número de muestras es pequeño. Siguiendo este principio, si todas las réplicas sometidas a un mismo tratamiento están segregadas (de cualquier forma) las diferencias encontradas podrían deberse al tratamiento o a la diferencia entre dos cosas diferentes de por sí.

Ej.: Estudiar las diferencias de mesofauna edáfica entre hayedos y encinares de la provincia de Madrid.

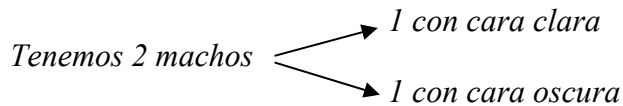


Según el objetivo planteado, tanto $x_1...x_6$ como $y_1...y_6$ se consideran pseudorréplicas porque son observaciones dependientes al compartir las mismas características ambientales. Si hubiera diferencias en la densidad de mesofauna entre ambos ecosistemas, no se sabe si es

[Alta] → 50 semilla (del mismo individuo adulto)

Estos datos serían pseudorréplicas porque las 50 semillas no son independientes al pertenecer al mismo individuo y las diferencias encontradas se pueden deber a diferencias genéticas.

Estímulos pseudorréplicados: ¿Prefieren las hembras de las cebras machos con cara clara o bien oscura?



Tenemos 20 hembras que se van colocando en el tiempo una a una y se anota la preferencia.

Resultado: 15 hembras prefieren el macho claro y 5 el oscuro.

Los estímulos (unidad experimental) son pseudorréplicas porque no están replicadas. Las hembras puede que elijan a los machos no por el color de la cara sino por otras características que no controlamos.

Solución: Habría que experimentar con varios machos claros y varios oscuros (Por ejemplo, a cada hembra se la expone a un par de machos de cada color).

¿Estaría bien el experimento si se introducen las 20 hembras simultáneamente en el recinto con los machos?

Solución: No porque las hembras pueden tener comportamiento social y no ser observaciones independientes.

Comparar el gasto mensual en dos barrios.

Serían pseudorréplicas si se seleccionan individuos de la misma familia o, quizás, individuos del mismo bloque de casas.

Un procedimiento para detectar la pseudorreplicación consiste en hacer un estudio pre-tratamiento en las unidades experimentales. Si el aumento del número de muestras aumenta la probabilidad de encontrar diferencias, se trata de pseudorréplicas. Si esto no ocurre tenemos auténticas réplicas.

En general, lo que se pretende es tratar de dilucidar cual es el grado de similitud para considerar dos (o más) unidades experimentales como réplicas. Esto es prácticamente imposible y además no es la cuestión, sino que lo que hay que hacer es replicar los tratamientos.

Las unidades experimentales heterogéneas producen una gran varianza del error experimental. La comparación precisa entre los tratamientos requiere la selección de unidades experimentales uniformes para reducir el error experimental. Sin embargo, una selección demasiado restringida puede producir condiciones de uniformidad artificiales. Un conjunto estrecho de condiciones restringe la base de inferencia del estudio. Entonces, para asegurar una confiabilidad razonable de las conclusiones del experimento, es deseable que las unidades representen una variedad suficiente de condiciones, sin que aumente innecesariamente la heterogeneidad de las unidades experimentales.

La naturaleza del experimento señala el equilibrio entre la variedad de las condiciones y la uniformidad de las unidades experimentales. *Por ejemplo, en un experimento con vacas lecheras (como es el estudiar el tipo de pasto que prefieren los adultos), si se desea una uniformidad de las unidades experimentales se requerirá que se elijan vacas de la misma especie y variedad, de la misma cría (criadas de la misma manera), en la misma etapa de*

vida, etc. Pero si se quiere inferir a cualquier vacase tendrá que perder uniformidad en las unidades experimentales.

Un caso grave por su difícil solución es el de los estudios de impacto ambiental, por la gran dificultad de replicar los tratamientos.

Evitar la pseudorréplica cuando se toman varias subunidades de muestreo dentro de una unidad de muestreo

Existen varias posibilidades dependiendo del interés del investigador:

- 1) **Media.** Esta opción es la más usada y nos ofrece el valor promedio de las observaciones.
- 2) **Mediana.** Ofrece un valor de la observación típica.

Ej: Yo tengo varios árboles y el número de frutos consumidos en cada uno por murciélagos.

- *Si calculo la media, tengo el número promedio de frutos consumidos por árbol. Representa el punto de vista del murciélago y de su actividad (índice de la cantidad de frutos consumidos por la población de murciélagos).*
- *Si calculo la mediana, tengo el número de frutos consumidos del árbol típico. Representa el punto de vista del árbol y sugiere la cantidad de frutos consumidos en ese árbol típico de la población, en ese año.*

- 3) **Varianza, Desviación típica o Coeficiente de variación.** Ofrece una medida de la variabilidad de los datos. La interpretación biológica de esta posibilidad es que ofrece una medida del grado de heterogeneidad-homogeneidad dentro de la unidad de muestreo.

Ej: En el caso puesto anteriormente de los hayedos y los pinares ¿Queremos saber cómo varía el grado de heterogeneidad en la abundancia de la fauna edáfica entre ambos sistemas?

¿Qué es más importante para su conservación, los cambios en la media o en la variabilidad de los datos?

- 4) **Sumar.** Siempre y cuando los valores de la variable sean sumativos.
- 5) **Alicuota.** También se pueden homogeneizar las submuestras en una sola y, a partir de ella, extraer una observación (alícuota) como valor representativo de la unidad de muestreo.

Pseudorreplicación temporal. Se produce cuando se toman diferentes unidades de muestreo en diferentes tiempos de una misma unidad experimental (es decir, por ej. múltiples medidas sobre el mismo individuo) y se consideran como independientes para los análisis estadísticos. *Por ejemplo, ver el efecto de la pluviosidad sobre la producción primaria de una especie de planta, para lo cual se estudia la producción en número de hojas sobre un mismo individuo durante varios años. O bien, número de frutos por individuo de esa especie durante varios años, usando siempre los mismos individuos.*

Tanto la pseudorreplicación simple como la temporal lo único que reflejan son variaciones dentro de la unidad de muestreo tanto en el espacio como en el tiempo, pero no son representativas del factor en que difieren las unidades de muestreo que comparamos. Por

ejemplo, si comparamos el número de especies de aves en un bosque pequeño y en un campo de cultivo próximo realizando varios censos en cada uno (bien simultáneamente o bien cada cierto tiempo), los resultados que obtenemos se refieren a la comparación entre ese bosque y ese campo, pero no a las diferencias generales entre bosques pequeños y campos de cultivo; para ello deberíamos censar muchos bosques y muchos cultivos similares, comparando los resultados conjuntos.

Ejemplos de Pseudorréplicación de medidas en el tiempo.

Ejemplo 1: Analizar las ventajas de dos insecticidas sobre los frutos de un mismo individuo en momentos diferentes (como es el caso de aplicar un insecticida en un año y el otro al año siguiente).

Ejemplo 2: ¿Prefiere el ciervo zonas abiertas (prado) o cerradas (bosque)?

Método. Se toman 15 momentos temporales, uno por día (12:00 h), durante 15 días.

- Se hace el seguimiento de un individuo. *Pseudorréplica*, ya que no es representativo.
- Se hace el seguimiento de varios individuos de una población. *Pseudorréplica*, ya que hay dependencia de los datos y no es representativo. Este tipo de poblaciones tiene estructura social y donde va el macho dominante va el resto del grupo.
- Se hace el seguimiento de varios individuos de diferentes poblaciones o de las poblaciones en sí como una única observación. **BIEN.**

Método. Se toman 15 momentos temporales a lo largo de un día (horas de luz, es decir desde las 7:00, ...hasta las 21:00 h). En este caso se puede incurrir en una pseudorréplica temporal si el hecho de que el ciervo esté en un tipo de zona está condicionado al punto temporal anterior y, a su vez, condiciona el dato siguiente. Es decir, entre un momento de muestreo y el siguiente no ha transcurrido el tiempo suficiente para que los ciervos puedan desplazarse entre las diferentes zonas o puedan recorrer todo el territorio, lo que hace que las observaciones no sean independientes. Esto puede ser también un ejemplo de lo que se denomina datos censurados, muy común en estudios de Etología (por ej. cuando la duración de un acto comportamental depende del tamaño de la unidad de muestreo en el tiempo).

También puede darse una pseudorreplicación filogenética cuando se da una compartición de un ancestro común, de manera que entre unas especies existe más independencia que entre otras. Por ejemplo, si queremos analizar como afecta la temperatura a la fisiología de los vertebrados de un ecosistema y se elijen al azar x vertebrados pero todos son mamíferos. Habría pseudorréplica porque todos ellos son filogenéticamente más dependientes que con los otros vertebrados (aves, reptiles, anfibios, etc.).

Pseudorreplicación sacrificada. Consiste en agrupar y testar muestras de verdaderas réplicas y, por lo tanto, considerar como independientes todas las observaciones de las diferentes unidades experimentales cuando, en realidad, no lo son.

También el agrupar supone implícitamente una redefinición de las unidades experimentales considerando a cada muestra como tal. Se pierde la información sobre la variabilidad de las réplicas.

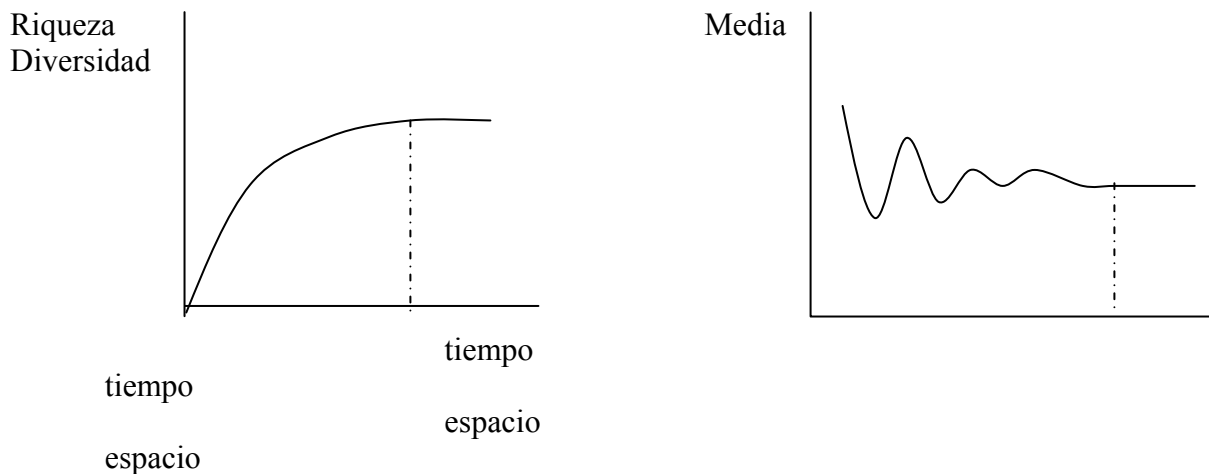
Por ejemplo, comparar las diferencias existentes entre las mirmecocenosis que caracterizan las partes altas y bajas de las laderas y para ello se agrupan por un lado los

datos de las partes altas de 4 laderas (20 observaciones en alto y 20 en bajo, representando cada una de ellas una edad sucesional) y por otro los datos de las partes bajas. Al agrupar estos datos se tienen un mayor número de observaciones por tratamiento, pero no serían verdaderas réplicas sino pseudorréplicas.

Poner como ejemplo bien hecho el de la práctica de *Lolium* en donde se agrupan los tiestos con diferente nivel de competencia para analizar el efecto de la diferente concentración de nitrato en el crecimiento de esta especie.

Determinación del tamaño de muestra

¿Cuál es el número de unidades de muestreo que debemos medir para obtener resultados representativos? es decir, ¿cuál es el tamaño de muestra mínimo? Es un hecho general que tanto la exactitud como la precisión de nuestros resultados aumentan a medida que aumenta el tamaño de muestra. Si medimos muy pocas muestras corremos el riesgo de obtener resultados equivocados; si medimos muchas, podemos estar perdiendo tiempo y esfuerzo de trabajo. Un procedimiento para ver la representatividad de nuestros resultados es **analizar como van variando a medida que aumentamos el esfuerzo de muestreo** (por ejemplo, representar el número acumulado de especies que se van encontrando en una zona a medida que aumentamos el tiempo de muestreo donde se puede ver que el número se incrementa al principio muy rápidamente, luego más lentamente y finalmente se estabiliza en un valor máximo a partir del cual no merece la pena seguir aumentando el esfuerzo de prospección).



De igual modo ocurre con los estadísticos (medidas que resumen nuestros resultados). Al principio fluctúan de forma más acusada, pero acaban estabilizándose a partir de un determinado tamaño de muestra. Estas relaciones pueden formalizarse matemáticamente, lo que permite calcular el tamaño de muestra requerido para alcanzar una determinada precisión en los resultados. En cualquier caso, las relaciones entre tamaño de muestra y precisión varían en cada caso particular, de modo que es necesario analizarlas en cada estudio a fin de poder calibrar correctamente los resultados obtenidos.

Poner el ejemplo del estudio de diversidad de hormigas de *Espadaler* y su jardín.

Uno de los métodos utilizados:

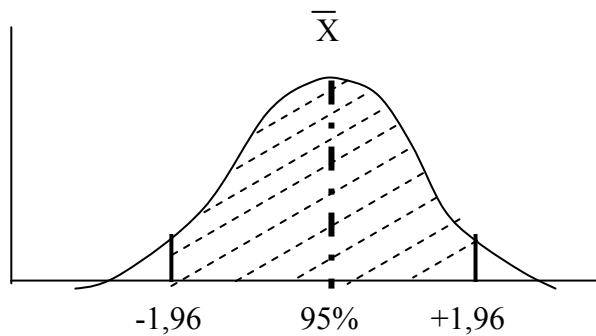
$$\sigma^2 (\square_{\alpha/2})^2$$

$$n = \frac{\sigma^2}{\varepsilon^2}$$

$\sigma^2 =$

$\alpha/2 =$ nivel de confianza elegido.

$\varepsilon =$ error que se quiere cometer (\pm) de la media muestral sobre la verdadera media poblacional.



Nivel de confianza (p)	Intervalo de confianza
1-α	Z $\alpha/2$
0,9	1,64
0,95	1,96
0,954	2
0,98	2,33
0,99	2,58
0,997	3

Ejemplo: Si se tiene una muestra aleatoria con $n=144$, una $\sigma^2=100$ \square $x=160$. Si se quiere tener un nivel de confianza del 95% de que la estima de la media se encuentra a una distancia de $\pm 1,2$ cm de la verdadera media poblacional, ¿cuántas observaciones adicionales deben tomarse?

$$n = \frac{100 (1,96)^2}{1,2^2} = 267$$

Según estos resultados el tamaño adicional de la muestra será

$$n = 267 - 144 = \mathbf{123 \text{ observaciones}}$$

Estadísticos básicos: medidas de tendencia general y medidas de variabilidad

Una vez seleccionada correctamente nuestra muestra y obtenidos los datos, lo primero que debemos hacer es resumirlos en una representación denominada **distribución de frecuencias**. A partir de ella se pueden ver gráficamente cuáles son los valores más frecuentes y cuál es la magnitud y la forma de la variabilidad en torno a ellos.

Comparando visualmente dos o más distribuciones uno se puede hacer una primera idea de hasta qué punto son diferentes, algo muy importante ya que casi todas las preguntas interesantes que podemos plantearnos van a suponer algún tipo de comparación. Pero lo

interesante y correcto es compararlas estadísticamente. Las pruebas que nos permiten comparar las distribuciones no se refieren normalmente al conjunto de los datos, sino que lo que se comparan son una serie de medidas que sirven para caracterizarlas. Estas medidas se denominan **parámetros** de la distribución cuando se refieren al valor real que alcanzarían si pudiésemos medir todo el universo muestral, y **estadísticos** cuando los valores se obtienen de nuestra propia muestra. Si la muestra está correctamente replicada, los estadísticos son medidas representativas (no sesgadas) de los parámetros y puede trabajarse con ellos con toda confianza.

No hay que confundir y distinguir claramente entre lo que es la inferencia estadística y la inferencia espacial y temporal antes comentada.

Los dos tipos de estadísticos más empleados son:

- Las medidas de tendencia central: **media, mediana** (en orden creciente, el valor central) y **moda** (valor de máxima frecuencia). Nos informan sobre los valores más frecuentes o representativos de la muestra y por tanto indican las características medias de nuestro universo muestral. Todas ellas son medidas de exactitud.
- Las medidas de variabilidad: **rango de valores, varianza, desviación típica y coeficiente de variación** ($C.V.=\sigma/x$). Nos informan sobre la precisión (dispersión) de los valores en torno a estos valores más frecuentes.

Con las actuales técnicas informáticas la obtención de estos estadísticos no plantea ningún problema y con sólo dos o tres números se pueden resumir completamente nuestros datos. Normalmente es suficiente con el valor medio, la varianza o la desviación típica y el número de unidades de muestreo que componen nuestra muestra para exponer nuestros datos.

Contraste de hipótesis

La base del contraste de hipótesis consiste en estimar la probabilidad de obtener al azar un nuevo valor determinado al comparar un dato con una muestra, dos estadísticos o dos series de datos (frecuencias o medidas simultáneas de dos variables). Es decir, a partir de los estadísticos que definen las distribuciones de frecuencias de nuestros datos, podemos establecer predicciones sobre la probabilidad de obtener datos adicionales si seguimos midiendo; estas predicciones se basan en el hecho de que los estadísticos son representativos del universo muestral (Por ejemplo, si se obtiene un valor medio de 17 en una muestra muy poco variable, lo más probable es que si medimos una nueva unidad obtengamos un valor próximo a 17). El hecho de que en estadística se trabaje con probabilidades, y no con certezas absolutas, implica que debemos adoptar algún tipo de criterio para aceptar o rechazar las hipótesis que planteamos en función de las probabilidades que obtenemos en un análisis estadístico. Este criterio se basa en establecer una probabilidad máxima a partir de la cual rechazamos la hipótesis de partida o **hipótesis nula** ($H_0 =$ nuestros resultados se deben al azar, es decir, los resultados se deben a la variabilidad natural y no hay ninguna causa real que los produzca). Si rechazamos la hipótesis nula, entonces aceptamos la **hipótesis alternativa** ($H_1 =$ los resultados se apartan de lo que cabría esperar por azar con lo que debe haber alguna causa que los produzca, es decir las diferencias se mantienen incluso teniendo en cuenta la variabilidad natural de los fenómenos estudiados).

Cuando tomamos la decisión de aceptar o rechazar la hipótesis nula, pueden darse cuatro situaciones:

- 1) Rechazar la H_0 cuando es realmente falsa (no hay problema).
- 2) Aceptar la H_0 cuando es realmente verdadera (no hay problema).
- 3) Rechazar la H_0 cuando es realmente verdadera (**error de tipo I** o α).
- 4) Aceptar la H_0 cuando es realmente falsa (**error de tipo II** o β).

Error de tipo I o α .-Las pruebas estadísticas disponibles calculan la probabilidad de cometer un error de este tipo y que es el más grave, ya que implica afirmar que existe un fenómeno donde no hay otra cosa que azar. Normalmente se asume como margen seguro una probabilidad del 95% o mayor de no equivocarse, con lo que α valdría un 5%. En tanto por uno, el criterio de rechazo de la hipótesis nula es de $\alpha < 0.05$. Cuanto menor sea la probabilidad de cometer un error de tipo I mayor es nuestra confianza de que los resultados obtenidos sean reales, no azarosos. Por ello se establecen niveles de significación del 99% ($\alpha < 0.01$, altamente significativo) y del 99.9% ($\alpha < 0.001$, muy altamente significativo) para expresar una mayor confianza en el resultado. El nivel de $\alpha = 0.01$ se establece para un solo test independiente. Ocurre que cuando se efectúan numerosas estimas de probabilidad se incrementa la probabilidad de que algunas sean significativas ($p < 0.05$) por puro azar. Si efectuamos N tests existirá una probabilidad > 0.05 de que alguno de ellos sea significativo por puro azar.

Por ejemplo ($p' = 1 - (1-p)^n$ en n tests):

1 test -----> probabilidad de cometer un error de tipo I=0.05

10 tests -----> probabilidad de cometer un error de tipo I=0.41

20 tests -----> probabilidad de cometer un error de tipo I=0.64

Esto es, al efectuar múltiples estimas de probabilidad (tests dentro de un mismo trabajo, tablas de correlaciones de distintas variables entre sí, etc.), se está incrementando la probabilidad de cometer el error de tipo I.

Lo más recomendable es hacer menos tests, sólo los estrictamente necesarios para examinar nuestras hipótesis.

Error de tipo II o β .- Mide la capacidad de una prueba de detectar diferencias significativas una vez eliminados los efectos del azar, es decir, rechazar la hipótesis alternativa cuando es verdadera; también suele denominarse **potencia de la prueba o potencia del test ($1-\beta$)**. Así, la potencia de un test es como hacer una lectura complementaria del error de tipo II, es decir, sería la probabilidad de rechazar la hipótesis nula cuando, de hecho, es falsa y la alternativa es la correcta; dicho con otras palabras, la potencia y probabilidad de cometer el error de tipo I (α) están inversamente relacionados. Cuando la hipótesis alternativa es general, es decir, cuando se limita a decir que debe haber alguna causa distinta del azar pero sin establecer cuál puede ser dicha causa, este error es máximo y se calcula como $1 - \alpha - \beta = \text{cte}$.

Para disminuir el valor de β es decir, para aumentar la potencia de un test existen varios métodos:

- Establecer hipótesis alternativas más concretas. ¿Por qué? Porque nos permite definir mejor el diseño. Es mejor saber mucho de poco, que poco de mucho.
- Aumentar el tamaño de muestra
- Emplear pruebas estadísticas paramétricas frente a pruebas no paramétricas siempre que esto sea posible.
- Si los efectos son pequeños se requieren muestras mayores.
- Si se quiere incrementa el nivel de significación hay que incrementar mucho los tamaños muestrales.

Las pruebas estadísticas se basan, como se ha dicho, en establecer la probabilidad de obtener un resultado determinado a partir de nuestros datos; si el cálculo de estas

probabilidades se basa en el empleo de la **distribución normal** se habla de **pruebas paramétricas**, que son, en general, más potentes a la hora de detectar efectos significativos. Si, por el contrario, nuestros datos no se ajustan a una distribución normal, existen dos opciones: o bien usar las denominadas **pruebas no paramétricas**, que no requieren que los datos se ajusten a ningún tipo concreto de distribución, o bien **transformar previamente los datos** hasta conseguir que se ajusten a una normal. En muchos casos, los datos transformados (calculando el logaritmo, el inverso, la raíz cuadrada, etc.) sí se ajustan a una distribución normal y, por lo tanto, pueden emplearse pruebas paramétricas, más potentes y que nos permiten establecer más conclusiones que las pruebas no paramétricas.

Las distribuciones se pueden caracterizar por dos parámetros que describen su forma:

- **Sesgo**. Mide la asimetría de la distribución respecto a lo que cabría esperar según la hipótesis nula de normalidad ($H_0=0$). Si es ≥ 0 está sesgada por la derecha; si es < 0 está sesgada por la izquierda.

- **Kurtosis**: Mide la puntiagudez de la distribución respecto a lo que cabría esperar según la hipótesis nula de normalidad ($H_0=0$). Si es ≥ 0 leptokúrtica; si es < 0 platikúrtica.

Consecuencias de la violación de la normalidad:

- Globalmente, los tests que utilizan la distribución de la F (Anovas) son bastante robustos a las desviaciones de la normalidad.

- El sesgo de las distribuciones no tiene un marcado efecto sobre la F (levísimo incremento del error de tipo I).

- La kurtosis tiene un efecto más marcado sobre el estadístico F.

Si $K > 0$, la F tiende a ser menor y corremos el riesgo de no poder rechazar la H_0 cuando ésta es incorrecta (incremento del error de tipo II).

Si $K < 0$, la F tiende a ser mayor, incrementando el error de tipo I.

Aplicación de pruebas estadísticas sencillas

Es muy importante tener en cuenta para emplear estas pruebas estadísticas:

- Planificar los diseños y las baterías de tests para poder ser lo más parsimonioso en nuestros análisis estadísticos.

- Establecer a priori los grupos de tests y las \square que serán utilizadas en los análisis.

- Hacer buena estadística equivale a conocer bien las técnicas y el concepto error.

- Hacer buena estadística supone hacer la menor cantidad de tests necesarios para examinar correctamente los efectos.

- Saber para qué tipo de hipótesis se aplican.

- Cuáles son los requerimientos mínimos en cuanto a los datos de partida.

- Cómo interpretar sus resultados.

Según el tipo de hipótesis a que se aplican y sus requerimientos en cuanto a la distribución de los datos se pueden clasificar:

A) **Pruebas de comparación de medias**. Prueban si las medias de dos o más muestras son estadísticamente significativas, planteando como hipótesis nula que las medias son idénticas.

A1) Dos muestras

• **Prueba paramétrica**. Exige distribución normal y tamaños de muestra grandes, en general mayores de 20: *t de Student*.

• **Prueba no paramétrica**. cualquier distribución o tamaños de muestra pequeños: *U de Mann-Whitney*.

A2) Más de dos muestras

- **Prueba paramétrica.** *Análisis de la varianza de clasificación simple (Anova).*
- **Prueba no paramétrica.** *Prueba de Kruskal-Wallis.*

Si la probabilidad que obtenemos es mayor de 0.05 aceptamos la hipótesis nula y concluimos que las medias no difieren significativamente; si es menor, concluimos que hay diferencias significativas y que las medias proceden, por tanto, de poblaciones diferentes.

B) Pruebas de comparación de frecuencias. Se aplica generalmente a la comparación de dos muestras, aunque pueden también usarse para comparar más de dos.

B1) *Test de bondad de ajuste.* Sirven para comprobar si una serie de frecuencias se ajusta a una distribución teórica dada, como por ejemplo la distribución normal.

B2) *Pruebas de contingencia.* Sirven para probar si dos o más series de frecuencias proceden o no de la misma distribución.

Las diferencias entre distribuciones se ajustan a la distribución χ^2 , con lo que se trata siempre de **pruebas no paramétricas**.

C) Pruebas de asociación. Prueban hasta qué punto los cambios en el valor de una variable están asociados (correlación) o determinan (regresión) los cambios de otra variable medida de manera simultánea en las mismas unidades de muestreo. La hipótesis nula es que no existe asociación y que ambas variables cambian de modo independiente.

- **Pruebas paramétricas.** *Correlación de Pearson y Regresión.*
- **Pruebas no paramétricas.** *Correlación de orden de Spearman.*

Si la probabilidad que obtenemos es mayor de 0.05 aceptamos la hipótesis nula y concluimos que las variables no están asociadas significativamente; si es menor, concluimos que ambas variables están asociadas. Si las distribuciones son normales, puede además medirse cómo la variación de una de las variables, que se asume cambia al azar (variable independiente) causa un cambio de la otra (variable dependiente), mediante un análisis de regresión, del que obtenemos una fórmula matemática (ecuación de la función) que nos permite calcular el valor de la variable dependiente para cualquier valor de la independiente.

D) Pruebas multivariantes. Se utilizan cuando se trabaja con un número elevado de factores, situación que no puede ser abordada por los análisis antes comentados.

D1) *MANOVA.* Análisis de la varianza multivariante. Muy sensible cuando hay valores fuera del rango normal.

D2) *Análisis de Componentes Principales, Análisis de Correspondencias, Análisis Discriminante, Cluster...* Son técnicas descriptivas. Pierden credibilidad cuando hay muchos ceros en la matriz. Los resultados cambian drásticamente cuando se quitan o se ponen valores “raros”.

Consideraciones sobre el diseño experimental

- Cuando se ejecuta el experimento es importante que el experimentador evite introducir errores sistemáticos (sesgo, variación sistemática de las mediciones por los aparatos de medida) y que minimice los errores aleatorios (ruido, variación aleatoria del método de medida y del material).
- El grado de heterogeneidad entre las unidades experimentales y el grado de regulación de las condiciones (ambientales) durante el experimento son subjetivos, no afectan a la validez formal del experimento.
- La función primaria de la estadística es ofrecer claridad, concisión y objetividad. Si hay errores estadísticos o de interpretación, los datos pueden ser analizados de nuevo, pero el único remedio para errores de diseño o ejecución (patrones, variación sistemática del material) es la repetición del experimento.

Las características principales (condiciones) de un experimento deben ser: controles, replicación, aleatorización y entremezcla.

- Controles. Son necesarios cuando la unidad de muestreo cambia con el tiempo. Por ejemplo, los sistemas biológicos.
- Replicación. La replicación aumenta la precisión de una estima. Reduce el efecto del "ruido" o variación aleatoria o error.
- Aleatorización. Elimina el posible sesgo del experimentador y garantiza que los errores están distribuidos aleatoriamente, que sean independientes estadística-mente (sólo de esta forma es posible conocer α , probabilidad de un error Tipo I).
- Entremezcla. Uno de los procedimientos de conseguir la entremezcla es mediante la aleatorización. Evita no apreciar gradientes o intrusiones no demoníacas (impregnación de cambios ocurridos sobre un experimento en proceso). Las demoníacas son las producidas por errores de diseño o ejecución.

POSIBLES PROBLEMAS

MEDIOS DE PREVENCIÓN/SOLUCIÓN

- Cambios con el tiempo -----> Controles
 - Efectos de procedimiento -----> Controles
 - Sesgo del experimentador -----> Aleatorización
(no independencia de las muestras)
 - Error aleatorio -----> Replicación
 - Variabilidad de las unidades experimentales -----> Replicación, Entremezcla
 - Intrusión de eventos casuales no detectados -----> Replicación, Entremezcla
- La entremezcla permite detectar efectos locales o accidentales al aumentar la varianza de los resultados dentro de cada tratamiento, ya que diferentes unidades del mismo tratamiento han sufrido diferentes intrusiones incontroladas.
 - La aleatorización es una forma de lograr la entremezcla (que no el fin) , de evitar que cada grupo de réplicas que recibe un tratamiento sufra una intrusión casual incontrolada diferente. Sin embargo, cuanto mayor es el número de réplicas mayor es la probabilidad de que se de (por azar) una segregación tal en las réplicas.
 - Otra forma de conseguir la entremezcla es gracias a diseños parcialmente aleatorios (agrupación en bloques, distribución sistemática, etc.), pero estos producen también, hasta cierto punto, una segregación.
 - La mejor forma de conseguir una buena entremezcla parece ser la propuesta por Cox (1958): Rechazar las distribuciones segregadas obtenidas al azar y realeatorizar hasta encontrar una con un grado aceptable (determinado de antemano) de entremezcla. Al estar los tratamientos mejor entremezclados se trabajará con un α menor que el inicial y, por tanto, la probabilidad de cometer un error de Tipo I será menor.

- "Los diseños sistemáticos pueden coincidir con un patrón sistemático de fluctuación del factor estudiado en el terreno" (Fisher). "Pero es más posible que coincida con un patrón correspondiente a una distribución aleatoria que a una regular, ya que las fluctuaciones en el terreno suelen ser irregulares y no periódicas" (Hurlbert).

¡ IMPORTANTE !

- En general, el problema es la representatividad de la (s) muestra (s) respecto al lugar (o tiempo) de donde ha (n) sido tomada (s).
- La diferencia entre unidades experimentales comparadas deberá ser mayor que las existentes entre diferentes puntos dentro de una unidad experimental, para decir que las unidades son diferentes.
- Si no hay réplicas o éstas no son estadísticamente independientes, el espacio real que estamos tratando está más restringido que el espacio inferencial que estamos suponiendo.
- Es importante tener claro qué es la unidad experimental y qué es el tratamiento.
 - Una forma de ver el efecto de cada uno de los parámetros sobre la unidad experimental sería por medio de controles que incluyen sólo uno de los elementos y excluyen los demás.

TIPOS DE ERRORES DE MUESTREO

Variación sistemática del material (errores de diseño o ejecución)

Si podemos distinguir un patrón en la variación entre individuos esto constituye una indicación de variación no aleatoria o sistemática. Por lo general, interpretaremos que ese patrón es debido a alguna de las escasas causas de efecto considerable si se compara con el de otras fuentes de variación, que consideramos como causas aleatorias. Por ejemplo, el peso en general más bajo de las plantas interiores (media 5,8 g) en una maceta que las plantas limítrofes (media 9,63 g) se ha interpretado como el efecto de competencia por la luz. Podríamos ignorar esto y tratar todas las plantas como pertenecientes a una misma población o, por el contrario, separarlas y tratarlas como dos poblaciones.

A veces la variación sistemática es más bien continua que discontinua. Por ejemplo, si hubiésemos cultivado rábanos en el suelo de un jardín con pendiente habría existido un gradiente de fertilidad de arriba a abajo, de modo que en conjunto las plantas tenderían a incrementar de tamaño en forma continua al ir descendiendo el nivel.

Si se dan estos errores lo mejor es repetir el experimento o entremezclar.

Variación aleatoria del método de las mediciones (ruido, errores aleatorios)

Si pesamos repetidamente una varilla de acero en una balanza, veríamos que los resultados sucesivos presentarían cierto grado de variación. De esta manera podríamos formar una población de pesadas en la que la variación fuese totalmente debida a la variación aleatoria del método de medición (ruido), ya que podría suponerse que la varilla mantendría virtualmente su peso constante. *Por ejemplo, la variación aleatoria en el peso fresco de las plantas de los rábanos, está formada por dos partes que nunca pueden ser totalmente separadas: a "ruido" y b variación aleatoria del material mismo.* Las mediciones deberían ser lo suficientemente precisas para que el ruido fuese muy inferior a la variación aleatoria del material, ya que sólo de esta forma puede estimarse esta última. *Por ejemplo, el método de medir la altura de los pinos con un clisímetro. Ej.: Semillas en cápsulas de jara (1.000 semillas); yo puedo contarlas todas (sin cometer sesgo) o contar 100 y a continuación hacer grupos de aproximadamente 100 y luego sacar el total (esto tiene sesgo).*

Variación sistemática de las mediciones (errores sistemáticos)

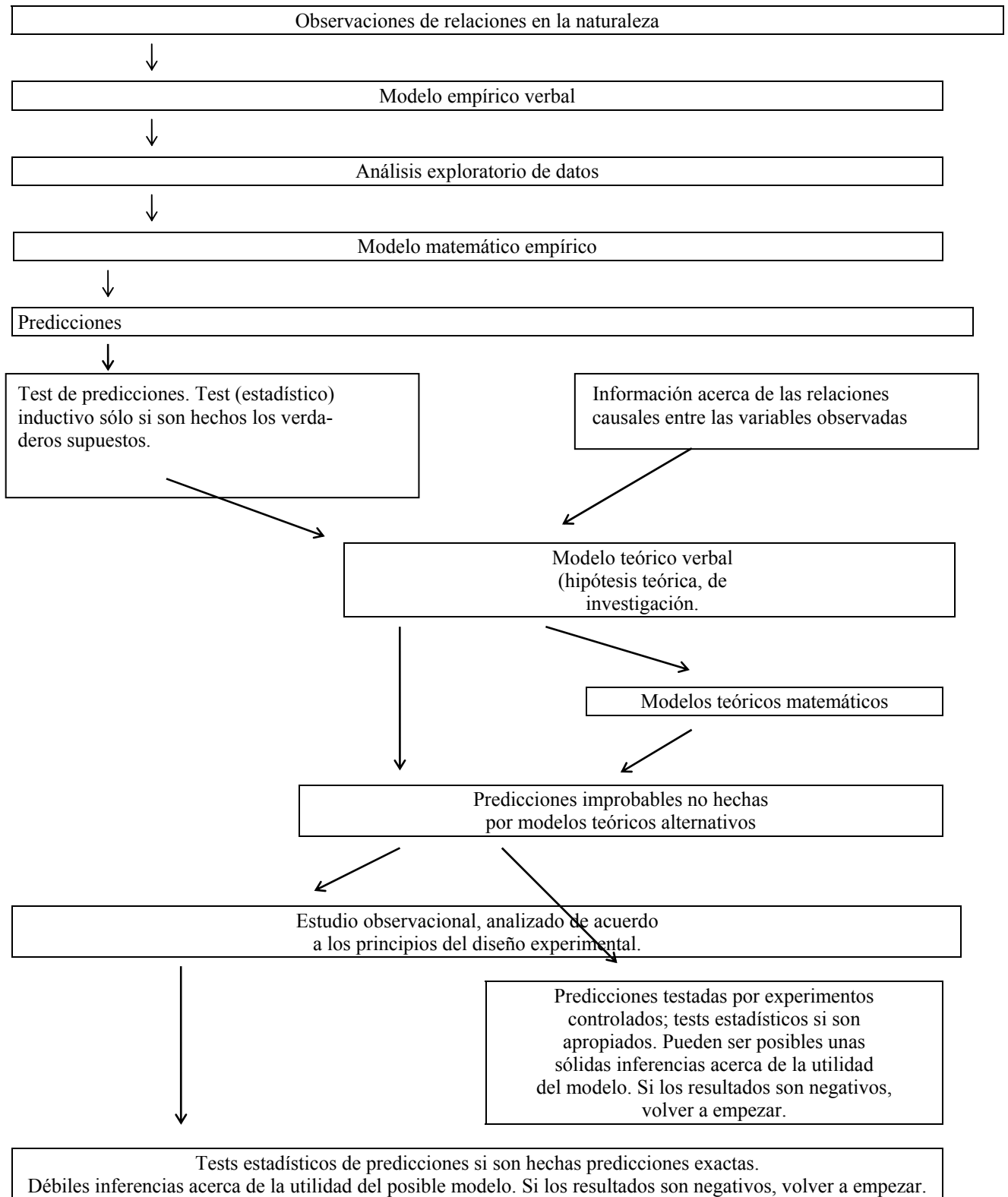
Esta variación hace referencia a los errores sistemáticos. Muchos instrumentos de medición dependen básicamente de cambios de resistencia eléctrica; estos cambios resultan a menudo sensibles a la temperatura y si no se efectúa la corrección adecuada el mismo valor de una variable puede dar distintas lecturas en distintas estaciones o en distintos momentos del día. El error sistemático no será importante si es muy pequeño en relación con la variación aleatoria del material y no haría falta perder tiempo y esfuerzo en hacer las correcciones necesarias para conseguir una medida perfecta.

ANÁLISIS DE DATOS

- ANÁLISIS DE CLUSTER.-** Métodos para construir diagramas ramificados (dendrogramas) que expresan la similitud (o diferencias) entre un conjunto de unidades (Pielou, 1977; Hartigan, 1975; Johnson & Wichern, 1982).
- ANÁLISIS DE COMPONENTES PRINCIPALES.-** Método estadístico multivariante diseñado para reducir la dimensionalidad de las variables originales creando una nueva composición de variables ordenadas por la varianza explicada (Johnson & Wichern, 1982, Tatsuoka, 1971).
- ANÁLISIS DE FUNCIÓN DISCRIMINANTE (DFA).-** Método estadístico multivariante para encontrar una nueva variable o variables (combinaciones lineales de las variables originales) que separa grupos previamente identificados en los datos (Pielou, 1977).
- ANCOVA.-** Una forma especial de ANOVA que permite ajustar variables independientes. Utilizado para incrementar la precisión en el análisis de varianzas al ajustar posibles sesgos introducidos por variables independientes y estimar medias de grupos ajustados para las variables independientes (Sokal y Rohlf, 1981; Snedecor y Cochran, 1980).
- ANOVA.-** Método numérico para repartir la suma total de cuadrados de los datos y sus grados de libertad. Frecuentemente se usa el test de la F para encontrar diferencias entre las medias de diferentes grupos (Sokal y Rohlf, 1981; Snedecor y Cochran, 1980).
- CLADÍSTICOS.-** En sistemática, grupo de métodos para construir diagramas ramificados (cladogramas) que expresa las relaciones de un conjunto de unidades (taxones) basados sobre la variación en rasgos discretos. Los métodos contienen criterios para determinar puntos de ramificación en los diagramas sobre la base de compartir caracteres derivados (Cracraft, 1982; Hull, 1979; Eldredge & Cracraft, 1980; Patterson, 1982).
- CORRELACIÓN CANÓNICA.-** Técnica estadística multivariante para calcular correlaciones entre una función lineal de un conjunto de variables y una función lineal de otro conjunto de variables. El objetivo es encontrar si hay correlaciones significativas entre los conjuntos (Sokal y Rohlf, 1981; Snedecor y Cochran, 1980).
- CORRELACIÓN MÚLTIPLE.-** Mide la asociación entre una única variable y una combinación lineal de un conjunto de variables (Sokal y Rohlf, 1981; Snedecor y Cochran, 1980).
- CORRELACIÓN PARCIAL.-** Mide la asociación entre dos variables mientras que ajusta para otras variables (Sokal y Rohlf, 1981; Snedecor y Cochran, 1980).
- CORRELACIÓN.-** Técnicas para medir la proximidad de relaciones lineales entre variables (Sokal y Rohlf, 1981; Snedecor y Cochran, 1980).
- ESTADÍSTICA NO PARAMÉTRICA.-** Métodos estadísticos que no asumen que los datos siguen una distribución particular (como es el caso de la distribución normal) (Conover, 1980).
- ESTADÍSTICA PARAMÉTRICA.-** Métodos estadísticos que dependen del hecho de asumir una distribución específica de los datos (por ej. normal) (Sokal y Rohlf, 1981; Snedecor y Cochran, 1980).
- HIPÓTESIS.-** Pueden ser hipótesis de investigación o hipótesis estadísticas. Las hipótesis de investigación son modelos teóricos que hacen predicciones. Las hipótesis estadísticas son exposiciones acerca de parámetros poblacionales. Normalmente, una hipótesis estadística, llamada hipótesis alternativa, refleja las predicciones de un modelo empírico o teórico y tiene un complemento lógico llamado hipótesis nula. El término hipótesis nula es un término estadístico y no necesariamente implica la situación descrita por esta hipótesis nula es un modelo nulo.
- JACKKNIFE.-** Técnica no paramétrica para encontrar un error estándar para estadísticos complicados (Sokal y Rohlf, 1981).

- MANOVA.- Análisis de la varianza multivariante, es decir, para más de una variable (Johnson & Wichern, 1982; Tatsnoka, 1971).
- MEDIA RECÍPROCA.- Técnica de ordenación indirecta (Gauch, 1982).
- MODELOS DE PROBABILIDAD.- Modelos que describen patrones de variación al azar. Modelos comunes: normal, binomial, Poisson, distribución t, chi cuadrado y distribución F (Sokal y Rohlf, 1981; Snedecor y Cochran, 1980).
- ORDENACIÓN.- Técnicas para ordenar unidades (por ejemplo, comunidades o especies) a lo largo de un eje o ejes. En la ordenación indirecta, los ejes están determinados por los patrones de co-ocurrencia de las especies. En la ordenación directa, los ejes están determinados por variables ambientales (Pielou, 1977; Gauch, 1982).
- PHENETICS.- Estudio de la similitud global de los fenotipos de grupos de organismos: métodos consistentes en reducir la dimensionalidad de los datos morfológicos o clasificar “unidades taxonómicas operacionales” (por ej. taxones) de acuerdo a la similaridad global (Sneath & Sokal, 1973).
- RAREFACCIÓN.- Método estadístico que puede ser utilizado para estimar el número de especies esperado en un muestreo al azar de individuos tomados de un censo (Simberloff, 1979; James & Rathbun, 1981).
- REGRESIÓN.- Técnica para modelar una variable (la variable dependiente) como una función de una o más variables (las variables independientes). Normalmente se utiliza para predecir los valores de la variable dependiente (Sokal y Rohlf, 1981; Snedecor y Cochran, 1980).
- TEST CHI-CUADRADO.- Diversos métodos estadísticos para testar, por ejemplo, test de bondad de ajuste y testar la independencia de variables categóricas (Sokal y Rohlf, 1981; Snedecor y Cochran, 1980).
- TEST DE LA t.- Test usado para comparar medias de distribuciones normales. También se utiliza como test de contraste en el análisis de la varianza y los coeficientes de regresión en las regresiones (Sokal y Rohlf, 1981; Snedecor y Cochran, 1980).

Progresión de pasos en el análisis de datos y el diseño experimental desde las observaciones a las conclusiones



TIPOS DE MODELOS

Modelos verbales. El proceso de formular modelos es probablemente la parte más creativa de la ciencia. Algunas veces sólo es posible establecer modelos verbales (no matemáticos) y son muy útiles. A niveles muy generales de la biología de las poblaciones, ecología de las comunidades, fisiología, biogeografía, morfología y evolución, los modelos verbales son útiles para organizar e interpretar la información. Un ejemplo de modelo verbal sería: “Los pájaros grandes tienen menores tasas metabólicas específicas/gr”.

Modelos empíricos. Son expresiones simbólicas de relaciones entre variables derivadas de los datos. Ejemplos de ellos son las relaciones (ratios), correlaciones ecuaciones de regresión, coeficientes de variación, índices de diversidad y similitud y ecuaciones multivariantes. Estos modelos sólo pueden ser utilizados para estudiar los resultados de los procesos. Las asunciones de un modelo empírico serían testadas para tratar de determinar si el modelo es útil en la práctica. Un ejemplo de modelo empírico sería la ecuación de regresión extraída de medidas en laboratorio que relaciona el metabolismo estándar con el peso del cuerpo en pájaros.

Modelos teóricos. Estos modelos podrían derivarse de un mejor entendimiento de los procesos. Su ventaja sobre los modelos empíricos es que pueden ser utilizados para estudiar procesos. Un buen modelo teórico simplifica los procesos estudiados reteniendo su esencia. Permite al investigador conceptualizar los procesos más fácilmente. Los modelos teóricos pueden ser verbales y matemáticos. Los modelos teóricos verbales no son otra cosa que modelos de procesos y su desarrollo matemático da lugar a los modelos teóricos matemáticos. Por ejemplo, los modelos teóricos matemáticos pueden ser manipulados para hacer predicciones fuera del campo en el que fue desarrollado. Ejemplos de modelos teóricos son: Modelo de equilibrio de biogeografía de islas de MacArthur y Wilson (verbal y matemático), Modelo de predador presa de Lotka-Volterra (modelo matemático), Teoría neo-Darwinista de la selección natural (no se demuestra con la experimentación, modelo verbal), Teoría del forrajeo óptimo (verbal y matemático).

Tanto con los modelos empíricos como con los teóricos se pueden hacer predicciones pero teniendo en cuenta lo siguiente. En el caso de los modelos empíricos se pueden hacer predicciones, si el sistema es el mismo que cuando fue formulado el modelo y las asociaciones utilizadas para derivar el modelo son fuertes. Sin embargo, si el sistema ha cambiado, o el modelo empírico es utilizado para hacer predicciones en un sistema completamente diferentes, éstas pueden ser muy pobres. Por el contrario, los modelos teóricos permiten hacer predicciones en cualquier caso.