

Índice

Capítulo 1. Experimentación y modelos estadísticos

<i>Susana B. Perelman y Lucas A. Garibaldi</i>	1
1.1. Incertidumbre.....	4
1.2. Preguntas, hipótesis, predicciones y objetivos.....	5
1.3. Diseño de experimentos.....	7
1.4. Tratamientos y unidades experimentales.....	9
1.5. Repeticiones y error experimental.....	11
1.6. Independencia entre unidades experimentales.....	14
1.7. Aleatorización.....	15
1.8. El número de repeticiones.....	16
1.9. Variabilidad: su control.....	18
1.10. Modelos estadísticos.....	20
1.11. Las pruebas de hipótesis estadísticas.....	25
1.12. Experimentos manipulativos y mensurativos.....	27
Ejercicios de aplicación	28

Capítulo 2. Análisis de la varianza y modelo

<i>Susana B. Perelman y Lucas A. Garibaldi</i>	31
2.1. Partición de la variabilidad.....	36
2.2. Partición de los grados de libertad.....	38
2.3. El cociente F y la prueba de hipótesis.....	40
2.4. El modelo lineal aditivo.....	44
2.5. Estimación de los parámetros del modelo.....	45
2.6. Esperanzas de los cuadrados medios.....	48
2.7. Comparaciones múltiples.....	50
2.8. El número de repeticiones.....	53
2.9. Intervalos de confianza en comparaciones múltiples.....	53
Ejercicios de aplicación	58

Capítulo 3. El control de la heterogeneidad

<i>Susana B. Perelman y Lucas A. Garibaldi</i>	67
3.1. Diseño completamente aleatorio (DCA).....	69
3.2. Diseño en bloques completos al azar (DBCA).....	72
3.3. Modelo para el DBCA.....	74
3.4. Partición de la variabilidad total en el DBCA.....	76
Ejercicios de aplicación	81

Capítulo 4. Validación de supuestos del ANOVA

<i>Susana B. Perelman y Karina Hodara</i>	85
4.1. Los residuales como elemento de diagnóstico.....	89
4.2. Independencia de los errores experimentales.....	89
4.3. Homogeneidad de varianzas entre tratamientos.....	90
4.4. Distribución de probabilidad normal de los errores experimentales.....	93
4.5. Transformaciones.....	100
Ejercicios de aplicación	102

Capítulo 5. Experimentos factoriales

Susana B. Perelman y Andrés G. Rolhauser 109

5.1. El arreglo factorial de los tratamientos 111

5.2. Modelo lineal para el experimento factorial completamente aleatorizado 119

5.3. Estimadores puntuales de los parámetros 120

5.4. Análisis de la varianza para un experimento factorial completamente aleatorizado . 121

5.5. Pruebas de hipótesis..... 124

5.6. Experimento con arreglo factorial y diseño en bloques..... 130

Ejercicios de aplicación 138

Capítulo 6. Análisis de regresión

Laura E. Puhl y Susana B. Perelman..... 143

6.1. Regresión lineal con una sola variable independiente 146

6.2. ANOVA para el análisis de regresión 157

6.3. Coeficiente de determinación R2 160

6.4. Regresión múltiple 162

6.5. Modelo general de regresión lineal múltiple 164

6.6. ANOVA para el análisis de regresión múltiple 167

6.7. El problema de la multicolinealidad 173

6.8. Regresión polinómica 178

6.9. Los supuestos en el análisis de regresión..... 181

Ejercicios de aplicación 196

Capítulo 7. Regresión con variable explicativa categórica

Fernando Biganzoli y Susana B. Perelman 195

7.1. El modelo, su interpretación y análisis 198

7.2. Análisis de covarianza (ANCOVA): supuesto de paralelismo 202

Ejercicios de aplicación 212

Capítulo 8. Bloques incompletos y parcelas divididas

María Virginia López 217

8.1. Diseño en Bloques incompletos balanceados (BIB) 220

8.2. Aleatorización en DBIB..... 222

8.3. Análisis de los DBIB 223

8.4. Variantes del diseño en bloques incompletos 230

8.5. Diseño en parcelas divididas..... 232

8.6. Ejemplo de experimento con diseño en parcelas divididas 239

8.7. Comentarios adicionales 242

Apendice A 243

Ejercicios de aplicación 246

Capítulo 9. Introducción a los modelos mixtos

Pedro M. Tognetti y Adriana Pérez..... 251

9.1. ¿Por qué son mixtos los modelos? Efectos fijos y aleatorios 254

9.2. El modelo mixto y sus parámetros..... 256

9.3. Modelos de regresión mixtos 260

9.4. Estructura de observaciones anidadas 267

9.5. Medidas repetidas en el tiempo 271

9.6. Varianzas heterogéneas 273

9.7. Consideraciones finales..... 277

Ejercicios de aplicación 279

Capítulo 10. Análisis de datos espaciales

Pablo A. Cipriotti 285

10.1. Los fenómenos con desarrollo espacial..... 288

10.2. Análisis de patrones de puntos 293

10.3. Análisis de variables continuas 300

10.4. Análisis de variables evaluadas en áreas 317

Referencias bibliográficas 323

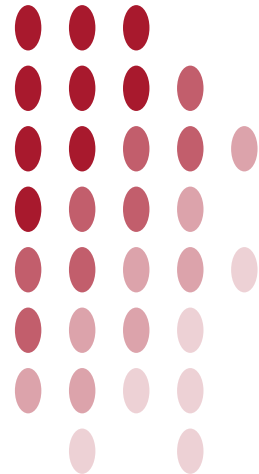
Capítulo 1

INTRODUCCIÓN A LA ESTADÍSTICA EXPERIMENTAL

Las motivaciones para experimentar

Los aspectos a tener en cuenta al planificar los experimentos

El modelo como síntesis de las preguntas del investigador



Introducción a la estadística experimental

Susana B. Perelman y Lucas A. Garibaldi

Los profesionales de las ciencias agropecuarias y ambientales enfrentan en su vida cotidiana múltiples situaciones en las que deben tomar decisiones, la mayoría de las veces basándose en información incompleta. Aquellos que se dedican a la investigación y participan en la generación de nuevos conocimientos, necesitan reglas para diseñar correctamente sus estudios. Otros, que no trabajan en investigación, se ven bombardeados frecuentemente por cúmulos de nueva información que deben seleccionar y evaluar objetivamente antes de incorporarla a su práctica profesional. Unos y otros requieren las herramientas que provee la estadística para establecer el grado de fiabilidad de los resultados generados en trabajos experimentales.

En la primera mitad del siglo XX los trabajos de R. A. Fisher (Figura 1.1) sentaron las bases del diseño de los experimentos agrícolas, proyectados específicamente para determinar los efectos de uno o varios factores sobre alguna variable de interés. Los mismos fundamentos se aplican actualmente para diseñar estudios que permiten determinar, por ejemplo, el efecto de la fecha de siembra y la aplicación de fertilizante nitrogenado sobre el rendimiento de un cultivo.



Figura 1.1: Izq. Sir Ronald Aylmer Fisher (1890 Londres-1962 Adelaida). Hizo enormes aportes a la Estadística Experimental y a la Genética Evolutiva. Imagen. Der. Foto aérea del experimento Park Grass de Rothamsted (<https://www.rothamsted.ac.uk/>), donde Ronald Fisher inició sus estudios del diseño de experimentos. Fuente: Wikipedia.

En la planificación de los experimentos, a través de una metodología para la colección de los datos y su análisis, se apunta a desentrañar el complejo de variables que intervienen en el fenómeno bajo estudio, apartando a todas aquellas que ocultan la acción de los factores que se desea estudiar. Los modelos estadísticos expresan a la variable de interés como combinación lineal de parámetros poblacionales asociados a los factores candidatos a explicar la variación de la respuesta. Mediante pruebas estadísticas referidas a los parámetros del modelo se pueden contrastar las hipótesis planteadas por el investigador y establecer restricciones a la interpretación arbitraria de la información obtenida.

1.1. Incertidumbre

Habitualmente en la vida profesional se deben tomar decisiones bajo incertidumbre. Por ejemplo, un evaluador ambiental debe decidir si la instalación de un gasoducto afecta significativamente la diversidad de especies de la selva que atraviesa, con la limitación que impone la imposibilidad de inventariar la diversidad de especies en TODA la superficie que puede ser afectada por el gasoducto y además sospechando fuertemente que los sitios no inventariados no tienen exactamente la MISMA diversidad de especies que aquellos evaluados. La estadística provee herramientas para cuantificar y disminuir la incertidumbre debida a estos dos aspectos fundamentales de la realidad: la variabilidad en las características estudiadas y la información incompleta.

Variabilidad en las características estudiadas. Uno de los rasgos más comunes de los fenómenos biológicos consiste en que sus manifestaciones no se repiten exactamente, de modo que cuando vamos a estudiarlas las mediciones que obtenemos nunca son iguales. Así, cinco plantas provenientes de semillas de girasol aparentemente idénticas, cultivadas en macetas con el mismo suelo y ambiente, muy difícilmente acumularán la misma biomasa o presentarán la misma altura a los 20 días de la emergencia.

Esta variación observada en los fenómenos biológicos se debe a múltiples causas. Por ejemplo, la biomasa de las cinco plantas de girasol podría diferir porque, entre otras razones, las 5 plantas tienen distinto genotipo, hay pequeñas diferencias en fertilidad o humedad del suelo entre las macetas, o la ubicación en el espacio no es la misma entre las 5 macetas. Generalmente los profesionales de las ciencias agropecuarias y ambientales están interesados en estudiar la variación en el fenómeno estudiado en función de alguna/s causa/s. Sin embargo, además de las causas que se estudian, los fenómenos biológicos están afectados por otros muchos factores que es necesario considerar a la hora

de interpretar los resultados obtenidos. Por ejemplo, supongamos el caso en que se desee estudiar el efecto del tipo de labranza (labranza convencional vs. siembra directa) sobre la invasión de malezas perennes en la Pampa Ondulada. La invasión de malezas podría verse afectada, no sólo por el tipo de labranza, sino también por los cultivos antecesores, por los cultivos actuales, por el tipo de suelos, por la distancia a fuentes de semillas, por la presencia de aves que las dispersen, o por el manejo diferencial de los cultivos que conducen los productores, entre otros muchos factores. En los siguientes capítulos se presentarán las herramientas que provee la estadística para disminuir la incertidumbre a partir de la cuantificación de la variabilidad.

Información incompleta. Usualmente no es factible observar todas las posibles manifestaciones del fenómeno estudiado, ya que no se pueden medir las variables de interés en todas las unidades de la población que se está investigando. Tal es el caso del evaluador ambiental que debe estudiar el impacto del gasoducto sobre la diversidad de la selva. Necesariamente debemos seleccionar un subconjunto de unidades muestrales sobre las cuales tomar los datos o realizar las observaciones y a partir de esa información incompleta extraer conclusiones acerca de la población.

1.2. Preguntas, hipótesis, predicciones y objetivos

La conducción de experimentos es una manera adecuada de responder preguntas acerca de la naturaleza, como las problemáticas planteadas en los ejemplos anteriores acerca de la sustentabilidad de los cultivos, de la diversidad de especies en la selva o de la invasión de malezas perennes en los cultivos. En muchos casos, los profesionales tienen posibles respuestas a estas preguntas antes de realizar el ensayo (i.e. **hipótesis**). Las hipótesis son afirmaciones sobre la realidad que solo pueden verificarse de manera indirecta, es decir, por alguna de sus consecuencias (i.e. **predicciones**). Entre las diversas formas que existen para generar conocimiento, el uso del método hipotético-deductivo ofrece la ventaja de permitir poner a prueba una idea que fue elaborada antes de realizar el estudio (hipótesis, Farji-Brener 2003). Este método consiste en deducir consecuencias concretas (i.e. generar predicciones derivadas de la hipótesis) sobre un determinado objeto de estudio bajo el supuesto de que nuestra afirmación o idea (hipótesis) es verdadera. De esta forma, al comparar los resultados obtenidos en nuestro ensayo contra los resultados que esperaríamos obtener si nuestra hipótesis fuese correcta, se puede poner a prueba la veracidad de la hipótesis y así avanzar en el conocimiento del problema planteado (Figura 1.2).

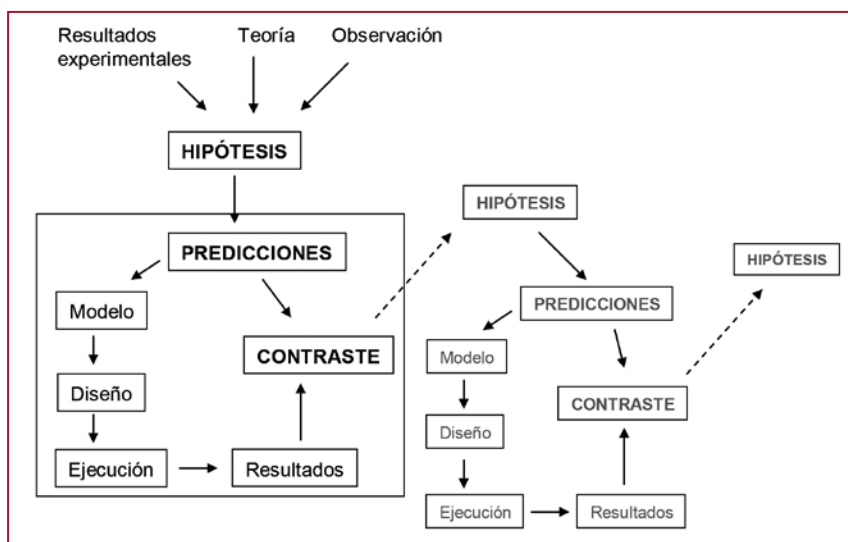


Figura 1.2: Esquema general de los componentes de un programa de investigación. El recuadro marca los límites del experimento. Del contraste de las predicciones de las hipótesis surgen conclusiones que generan nuevas hipótesis

Por ejemplo, si se postula que la siembra directa (SD) en la Pampa Ondulada favorece la invasión de los campos con malezas perennes, la predicción derivada de esta hipótesis sería que en lotes con larga historia de siembra directa se encontrará una mayor abundancia de malezas perennes que en aquellos lotes que durante el mismo tiempo se cultivaron con labranza convencional (LC). Un experimento correctamente diseñado permitiría verificar si esta predicción se sostiene en los lotes cultivados de la Pampa Ondulada. Para ello se deberán contestar varias preguntas de este tipo:

- ¿Cómo se va a medir el efecto de la SD sobre la invasión con malezas perennes?
- Además del tipo de labranza (SD vs. LC), ¿qué otros factores podrían afectar las respuestas que se van a analizar?
- ¿Cuáles de esos factores se van a considerar en el estudio?
- ¿Cuál es la población a la que se van a referir las conclusiones?
- ¿Cuál sería la forma de analizar los datos?
- ¿Qué diferencias en la abundancia de malezas se van a considerar importantes?

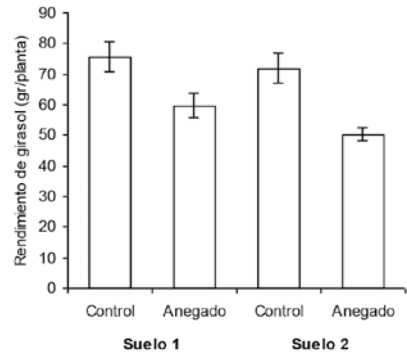
- ¿Cuántas veces se deberían repetir las observaciones para obtener resultados confiables?

El **objetivo** enuncia qué es lo que se pondrá a prueba en la investigación y cuál es el objeto de estudio. En el ejemplo de la invasión de malezas, el objetivo del trabajo podría ser: Estudiar el efecto del tipo de labranza sobre la invasión de malezas perennes en lotes cultivados de la Pampa Ondulada.

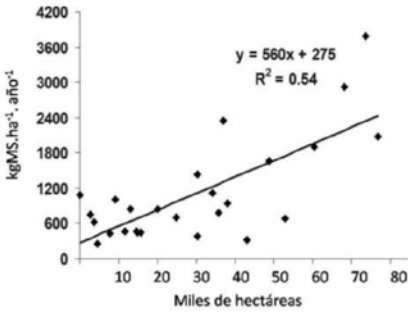
1.3. Diseño de experimentos

Realizar un buen diseño es planificar el experimento de manera que se pueda obtener la información adecuada al menor costo posible. Es decir, asegurar que se tomarán los datos que conduzcan a decisiones válidas en relación al problema planteado, minimizando los costos en tiempo y esfuerzo. Las reglas contenidas en los principios básicos del diseño experimental ayudan a generar conocimientos de manera objetiva, precisa, sistemática y repetible. La planificación del experimento es una etapa crítica e irreversible en este proceso, ya que ningún análisis posterior puede convertir un mal diseño en uno bueno. Si el experimento no es correctamente diseñado, no importa cuán bueno sea el análisis posterior de los datos, ya que no se podrá cumplir con el objetivo del estudio (Eberhardt & Thomas 1991).

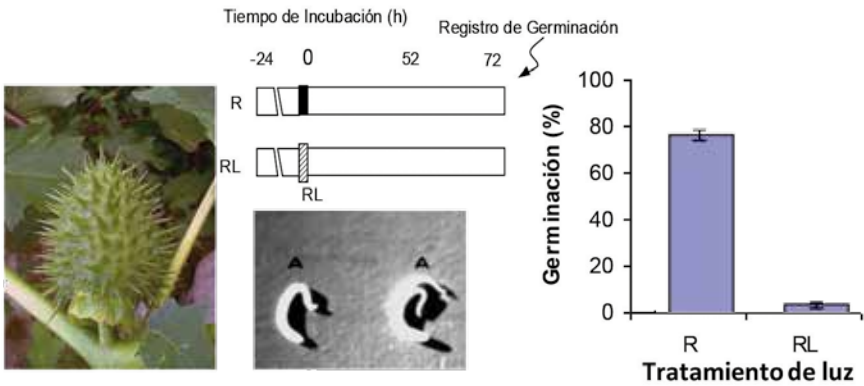
Los experimentos conducidos en laboratorio, en invernáculos o a campo, no difieren en su estructura lógica; así sea que se trate de estudiar el efecto de la calidad de la luz sobre la germinación de semillas de chamico en cajas de Petri, el crecimiento de plantas de girasol en macetas con distinta fertilidad de suelos o la relación entre el área desmontada y el consumo de biomasa por fuego en el bosque Chaqueño (Figura 1.3). A pesar de las diferencias de contexto, todos estos estudios encuentran dificultades semejantes para aislar los efectos buscados, porque la interpretación de cualquier fenómeno biológico requiere prestar atención a los obstáculos antes mencionados: la variabilidad en las respuestas, las causas múltiples que las afectan y la incapacidad de observar la población completa. En la consideración de estos aspectos se fundamentan los principios básicos del diseño de los experimentos: la *repetición*, la *aleatorización* y el *control de la heterogeneidad*, que serán discutidos a continuación (secciones 5 a 9). Previamente presentaremos la definición de *tratamientos*, *unidades experimentales* y *variable respuesta*, elementos indispensables para esa discusión.



1.3a: Efecto del anegamiento sobre el rendimiento de plantas de girasol en macetas con distintos suelos



1.3b: Relación entre el consumo de biomasa por fuego (kgMS.ha-1. año-1) y el área desmontada para los departamentos de la provincia de Santiago del Estero (período 2002-2006). (Adaptado de Rueda *et al.* 2012)



1.3c: Germinación de semillas de *Datura ferox* (chamico) en respuesta a la calidad de la luz

Figura 1.3: Experimentos conducidos en diferentes escalas. Identificar las unidades experimentales en cada uno

1.4. Tratamientos y unidades experimentales

Se denomina **tratamientos** a los diferentes procesos cuyos efectos van a ser medidos y comparados. Por ejemplo, podríamos estar interesados en comparar el rendimiento de cuatro híbridos de maíz, o el efecto de tres niveles de riego (sequía, inundación y riego normal) sobre el crecimiento de un cultivo, o el efecto del cultivo con diferentes especies remediadoras sobre la concentración de plomo en un suelo contaminado, o la efectividad de distintas vacunas para el control de la aftosa. En todos estos casos el experimento consiste en aplicar a cada *unidad experimental* (parcela, maceta, animal, establecimiento, planta, etc) uno de los distintos tratamientos a comparar y medir la *respuesta* que producen, para poder luego estimar sus efectos.

Una **unidad experimental** es la mínima porción de material experimental que puede recibir un tratamiento (y que, en consecuencia, podría haber recibido cualquier otro) y sobre la cual se observará una respuesta para evaluar el efecto del tratamiento. En los ejemplos de la Figura 1.3 las unidades experimentales son respectivamente: cada maceta con una planta de girasol, cada departamento de la provincia de Santiago del Estero y cada caja de Petri con 20 semillas de chamico. Las unidades experimentales pueden tener distinto tamaño y forma, dependiendo por ejemplo de la especie en estudio de la heterogeneidad ambiental presente o del tipo de tratamiento a asignar.

Las preguntas que intentamos responder mediante un experimento o un plan de investigación se refieren generalmente a una población, que debe ser explícitamente identificada al momento de planificar el estudio. Habitualmente es improbable que podamos observar o estudiar todas las unidades experimentales de la población, por lo tanto, se obtiene una **muestra** representativa de la población. Es decir, se selecciona un *subconjunto* de unidades experimentales (muestra) del *conjunto* de todas las unidades experimentales (población). Por ejemplo, si estamos interesados en estudiar el diámetro de los troncos de los árboles de álamo cultivados en el delta del Río Paraná (población de referencia), la unidad experimental será cada árbol. Una muestra podría ser un conjunto de 150 árboles de álamo seleccionados al azar en las explotaciones comerciales del delta del Paraná. Como se mencionó anteriormente, habitualmente nos manejamos con información incompleta ya que no contamos con mediciones realizadas en todas las unidades que constituyen la población. Es decir que seleccionamos una muestra para tomar decisiones o realizar conjeturas, métodos estadísticos mediante, respecto de la población completa.

Finalmente, es muy importante que las unidades experimentales sean representativas de la población sobre la que se han fijado los objetivos del estudio. Por ejemplo, en relación con los pastizales inundables del SE de la provincia de Buenos Aires, si se realizaran experimentos de manejo de pasturas en una estación experimental situada en Balcarce, con suelos profundos y muy ricos en materia orgánica, no correspondería extrapolar los resultados para asesorar a los productores del partido de Castelli, quienes manejan pasturas en ambientes con problemas de inundación, aunque la distancia geográfica entre ambos es menor a 200km.

Cualquier pregunta que queramos responder mediante un experimento implica realizar alguna observación o medición (**variable respuesta**) sobre cada una de las unidades experimentales de la muestra. De este modo, en el ejemplo de los álamos, podríamos estar interesados en observar la altura, el diámetro a la altura del pecho, el volumen del árbol o la rectitud del fuste entre otros, dependiendo del objetivo del estudio. La variable observada o de respuesta puede ser **cuantitativa** o **cualitativa**. Una variable cuantitativa asume valores numéricos que poseen un orden inherente, tal como los diámetros de tallos de plantas o las edades de un grupo de empleados. A su vez, una variable cuantitativa puede ser **cuantitativa discreta** (número de hojas por planta, cantidad de pétalos por flor) o **cuantitativa continua** (diámetros de tallos, pesos de semillas). Una variable cualitativa o categórica, es aquella cuyos valores corresponden a una u otra categoría de un conjunto de categorías que son mutuamente excluyentes. Por ejemplo, infectada o no infectada son dos posibles categorías de la variable plantas examinadas por su sanidad. Igualmente, rojo, azul y blanco podrían ser tres posibles categorías para la variable color de los pétalos.

Ocasionalmente los tratamientos no pueden ser aplicados o manipulados por el investigador, sino que ya vienen incluidos en las unidades experimentales (Hulbert 1984). Por ejemplo, si se desea comparar la facilidad de incorporación de nuevas prácticas agropecuarias entre productores con distinto nivel de educación es obvio que los tratamientos (educación primaria, secundaria y universitaria, por ejemplo) no pueden ser aplicados a los productores por el investigador. De igual forma, si se quiere estudiar el crecimiento de árboles de lenga en relación con la exposición (norte o sur) de las laderas en las que se encuentran los bosques, el factor exposición no puede ser aplicado, sino que las unidades experimentales ya están asignadas a un determinado tratamiento por su ubicación espacial. En estos casos, los tratamientos no se manipulan, sino que simplemente se registran sus valores; se aprovecha la variabilidad natural en las características de interés para ver si se ajustan o no a lo que

predicen las hipótesis que se formularon. Algunos autores denominan a este tipo de tratamientos **factores de clasificación** y a los experimentos que los estudian **mensurativos** en contraposición a los **experimentos manipulativos**, en los cuales el investigador aplica los tratamientos. La distinción no tiene relación con la precisión de los instrumentos de medida utilizados en uno y otro caso, ni con el lugar de la toma de datos (campo o laboratorio), sino sólo con el hecho de que se pueda aplicar o no un tratamiento a cualquier unidad experimental (Figura 1.4).

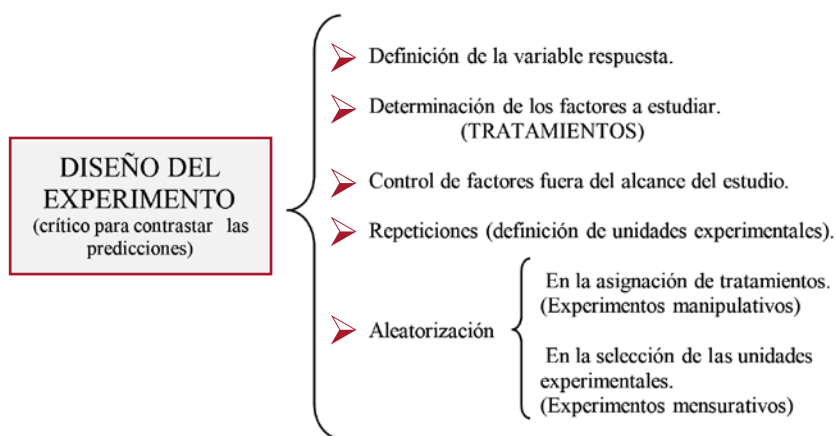


Figura 1.4: Principales elementos a considerar en el diseño experimental

1.5. Repeticiones y error experimental

La acción de repetir consiste en asignar cada uno de los tratamientos a varias unidades experimentales independientes. Dado que ningún animal, planta o parcela es idéntico a otro, siempre habrá alguna variación en las respuestas que no ha sido controlada, aún cuando el experimento sea conducido muy cuidadosamente. En consecuencia, los resultados de los experimentos están afectados no sólo por la acción de los tratamientos que aplicamos, sino también por las causas múltiples no controladas que tienden a encubrir sus efectos (Figura 1.5). Estas diferencias en la respuesta entre unidades que recibieron el mismo tratamiento, que en consecuencia son diferencias no debidas al efecto de los tratamientos, se denominan errores experimentales (Oddi *et al.* 2018). En otras palabras, el error experimental es una medida del fracaso en alcanzar resultados idénticos en unidades experimentales que fueron tratadas idénticamente.

La repetición de unidades experimentales independientes expuestas al mismo tratamiento permite estimar el error experimental y utilizarlo como unidad básica de medida para evaluar el significado de las diferencias observadas entre tratamientos. “Error” no significa que nos hayamos confundido, el error experimental representa el efecto del resto de los factores (que no son de interés para el estudio) sobre la variable respuesta (Oddi *et al.* 2018).

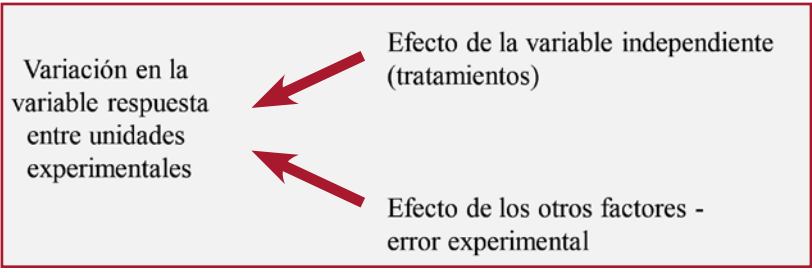


Figura 1.5: Determinantes de la variación entre distintas unidades en la variable respuesta

Por ejemplo, si se desea comparar el rendimiento de dos híbridos de maíz (híbridos A y B), 5 repeticiones por tratamiento implicaría asignar 5 parcelas a cada híbrido, de modo que el experimento contaría con 10 unidades experimentales en total (Figura 1.6). Como podemos observar, existe variación en el rendimiento entre las unidades experimentales, es decir, las 10 unidades experimentales no presentan el mismo rendimiento. Una parte de esa variación se debe al efecto de los tratamientos, o sea, a las diferencias entre los hí-

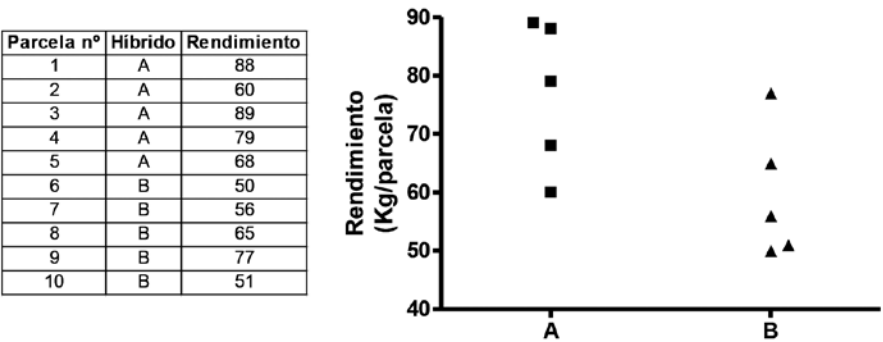


Figura 1.6: El rendimiento observado en cada parcela varía según el híbrido sembrado (tratamientos) y según otros factores (error experimental). A= Híbrido A, B= Híbrido B

bridos. La otra parte de esa variación se debe al error experimental, es decir, a diferencias entre las parcelas en cuanto a todos los otros factores que afectan al rendimiento (topografía, historia de uso, manejo, etc.). Por lo tanto, el error experimental abarcaría las diferencias en rendimiento observadas entre las parcelas que fueron cultivadas con el mismo híbrido. De este modo, si observamos dos parcelas con el híbrido A, el mayor rendimiento observado en la parcela 3 con respecto a la 4 no se debe a diferencias entre los híbridos. En cambio, este mayor rendimiento si se debe a otros factores, por ejemplo, tal vez la parcela 3 tuvo menor historia de uso o mayor humedad que la parcela 4.

Cuando los tratamientos no se encuentran repetidos en varias unidades experimentales, es imposible discernir si las diferencias observadas se deben a diferencias entre tratamientos o a diferencias intrínsecas entre las unidades experimentales. Esto nos impide evaluar de manera concluyente el efecto de los tratamientos. Por ejemplo, supongamos el caso en que no existen diferencias entre esos dos híbridos, y realizamos un ensayo con 5 repeticiones. Los resultados podrían tener un aspecto similar a los de la Figura 1.7a. En el caso en que hubiéramos realizado este experimento con una sola repetición, es decir una parcela con el híbrido A y otra parcela con el híbrido B, podríamos haber observado algo similar a la Figura 1.7b ya que nunca dos unidades experimentales son iguales. La diferencia en el rendimiento entre esas dos parcelas, puede deberse a una diferencia desconocida de fertilidad de suelo entre los lotes o a una genuina diferencia de potencial de rendimiento entre los híbridos. El agrónomo realizando el ensayo interesado en detectar diferencias entre los híbridos podría sugerir incorrectamente que el híbrido B tiene un mayor rendimiento que el híbrido A. Obviamente, el agrónomo que conduce un estudio no está interesado en las diferencias de rendimiento entre los dos híbridos en *esos* dos lotes en particular: él desea que sus resultados puedan ser aplicados a muchos otros lotes semejantes que no participan de su ensayo. En consecuencia, necesita alguna manera de discernir si las diferencias observadas se encuentran dentro de los límites del error experimental o si, por el contrario, son suficientemente grandes como para no poder ser explicadas meramente por azar. O sea, cuando no existe un método para estimar el error experimental, no hay manera de determinar si las diferencias observadas indican un efecto puro de los tratamientos o si se deben a la variación inherente en el material al que fueron aplicados. En este caso, se dice que hay **efectos confundidos** ya que no se pueden separar los efectos de los tratamientos de los efectos de otros factores que no son de interés en el estudio.

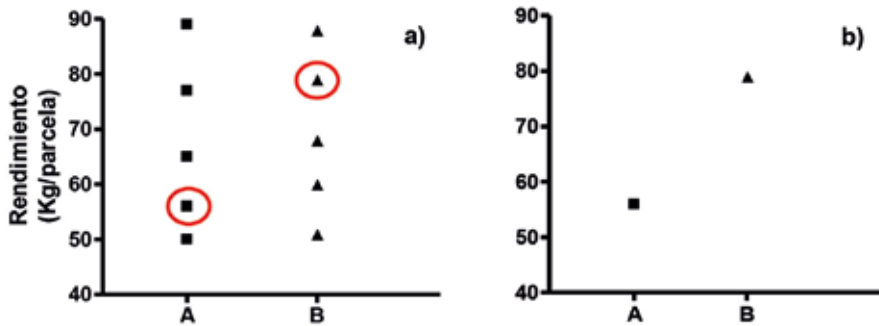


Figura 1.7: Cuando los tratamientos no son repetidos en varias unidades experimentales, no se puede discernir si las diferencias observadas se deben a diferencias entre los tratamientos o a diferencias intrínsecas entre las unidades experimentales. A=Híbrido A; B=Híbrido B

1.6. Independencia entre unidades experimentales

En el contexto del diseño de experimentos se entiende por independencia entre unidades experimentales que el valor observado para la variable respuesta en una unidad experimental no sea afectado o no esté relacionado con los valores de otras unidades experimentales. Nos interesa que:

- El efecto de los tratamientos sea independiente del efecto de otros factores (i.e. las unidades experimentales asignadas a un mismo tratamiento deben ser independientes entre sí).
- Las unidades experimentales asignadas a distintos tratamientos sean independientes entre sí (i.e. los tratamientos deben ser independientes entre sí).

Por ejemplo, supóngase que se desea probar el efecto de un herbicida de contacto aplicado por pulverización (dos tratamientos, con y sin herbicida) sobre el porcentaje de cobertura de una maleza en lotes sembrados con un cultivo (la unidad experimental es cada lote). En el primer caso, una falta de independencia implicaría que alguno de los tratamientos se vea favorecido o perjudicado en la asignación de unidades experimentales. Por ejemplo, que los lotes o parcelas tratados con el herbicida presenten una menor historia de uso y por lo tanto una menor cobertura de la maleza previa a la aplicación del herbicida, que aquellos lotes a los cuales no se les aplicará el herbicida. En el segundo caso, una falta de independencia podría darse, por ejemplo, si las parcelas estuvieran muy cerca unas de otras y el herbicida se aplica por fumigación en un día con

alta velocidad de viento. En este caso, el herbicida llegaría a las parcelas que no deberían recibirlo debido a la deriva provocada por el viento. De este modo, las acciones tomadas en las unidades experimentales correspondientes al tratamiento con herbicida afectaron los resultados de las unidades experimentales correspondientes al tratamiento sin herbicida. Por lo tanto, no hay independencia entre las unidades experimentales asignadas a distintos tratamientos. El investigador debe realizar al menos dos tareas para favorecer la independencia entre unidades experimentales, ellas son la aleatorización y el control del experimento (i.e. que las parcelas “testigo” no reciban el herbicida debido a la deriva provocada por el viento).

1.7. Aleatorización

Para evitar el sesgo en las comparaciones entre medias de tratamientos, es necesario asegurar que un tratamiento particular no resulte favorecido en forma consistente en repeticiones sucesivas por alguna fuente desconocida de variación. Por lo tanto, un supuesto importante de cualquier experimento, es que las unidades experimentales no presentan diferencias entre los tratamientos anteriores a la aplicación de los mismos. Es decir, las **condiciones iniciales** son las mismas para las unidades experimentales que reciben los distintos tratamientos.

En los experimentos manipulativos, *cada tratamiento debe tener igual oportunidad de ser asignado a cualquier unidad experimental*. Al asignar al azar los tratamientos a las unidades se obtiene un procedimiento que logra, en promedio, una distribución equitativa de los efectos de múltiples factores desconocidos. De esta forma, la aleatorización, combinada con la repetición, son requisitos para obtener una medida válida del error experimental. La aleatorización permite que las decisiones acerca de la asignación de las unidades experimentales a diversos tratamientos no sean el producto de alguna distorsión, regla o parcialidad arbitraria y sistemática. La aleatorización evita que una arbitrariedad seguida sistemáticamente por el investigador o por alguien más, haga tendenciosos los resultados.

En los experimentos mensurativos, donde los tratamientos son inherentes al material, la aleatorización interviene en la selección de las unidades experimentales sobre las que se toman los datos u observaciones. El muestreo es necesario porque comúnmente no se pueden medir las variables en todas las unidades de las poblaciones que estamos comparando. De esto se deduce de forma inmediata que un buen muestreo es aquel en el que la muestra es representativa del fenómeno o de la población estudiada, de manera tal que

los resultados obtenidos se puedan extender a otras situaciones de la misma población que no van a ser medidas. El mejor método para garantizar la independencia de las unidades experimentales en lo que respecta a factores no controlados, es seleccionar al azar, de entre todas las posibles, las unidades que se van a medir.

La aleatorización es imprescindible pero no suficiente para alcanzar la independencia. Además de aleatorizar la asignación de los tratamientos a las unidades, el experimentador debe realizar todas las acciones que considere necesarias para favorecer la independencia entre las unidades experimentales. Retornando al ejemplo del herbicida de contacto, es importante que la aplicación por fumigación, así como cualquier acción desarrollada en las parcelas con el tratamiento herbicida, no afecte los resultados de las parcelas que no recibirán el herbicida (la recíproca es igualmente válida). En este caso, el experimentador debe aplicar el herbicida de manera cuidadosa, en un día con baja velocidad de viento y las parcelas deben estar lo suficientemente alejadas entre sí. En general, un buen control experimental implica una adecuada separación espacial y temporal entre las unidades experimentales.

1.8. El número de repeticiones

¿Qué es lo que puede decirse acerca de la diferencia verdadera entre dos tratamientos? El análisis de los resultados no puede dar el valor exacto de la diferencia verdadera (que no se conoce), y tampoco es posible determinar dos límites dentro de los cuales tengamos certeza de encontrar el valor exacto. En cambio, se puede hacer una afirmación para una probabilidad elegida, por ejemplo 0.95 o 0.99, de que la diferencia verdadera se encuentra entre dos límites calculados con los datos. Es decir, se puede establecer la probabilidad de que sea correcta la afirmación que establece límites para la diferencia verdadera.

¿Será esta información suficientemente precisa para que haga posible tomar decisiones de importancia práctica? En general, sí, a menos que los datos hayan sido insuficientes en cantidad o en calidad; por ejemplo, supongamos que la aplicación de una nueva tecnología de cultivo resultara beneficiosa sólo en el caso de que aumente el rendimiento en 80 kg (porque en ese caso el beneficio supera al costo diferencial de aplicar la nueva práctica). Supongamos además que se realiza un experimento y se obtienen límites de confianza para el efecto de la práctica de 130 y 345 kg; entonces se puede recomendar su aplicación con la tranquilidad de que resultará efectiva. De igual forma, si los límites resultaran -15 y 45 kg, se decidiría no introducir la nueva práctica. En cambio,

si los límites resultaran de 10 y 165 kg, no se puede hacer ninguna recomendación sin gran riesgo de equivocarse. Para tomar la decisión correcta en esta última situación se debe realizar un experimento que acorte la distancia entre los límites de confianza.

La herramienta más conveniente que tiene el experimentador para dar a sus ensayos mayor capacidad de detectar diferencias más sutiles entre tratamientos es incrementar el **número de repeticiones**. Un experimento realizado con repeticiones independientes, pero con pocas repeticiones, puede resultar inútil para solucionar un determinado problema. Por tradición, en los experimentos agrícolas se han empleado tres o cuatro repeticiones sin considerar la importante influencia que tiene esta decisión sobre la capacidad del experimento para detectar diferencias significativas entre los tratamientos. Un experimento con pocas repeticiones solo será capaz de detectar diferencias muy groseras, sin embargo al incrementar el número de repeticiones aumenta simultáneamente el esfuerzo de la experimentación, de manera que resulta de especial importancia encontrar el número conveniente de repeticiones para cada experimento. Por ejemplo, supóngase que se desea evaluar el efecto del raleo (remoción de árboles) sobre el incremento en diámetro de árboles de lenga (*Nothofagus pumilio*) en bosques naturales. Para ello, se seleccionarán parcelas de bosque natural, a las que se les aplicará al azar una de tres prácticas de manejo: a) sin raleo, b) raleo bajo y c) raleo alto. ¿Cuántas repeticiones deben realizarse? Es decir, ¿cuántas parcelas deben ser asignadas a cada uno de los tratamientos? La respuesta a esta pregunta depende de las características de la población con la que se está trabajando así como de los intereses/objetivos del investigador. A continuación se discutirán conceptualmente algunos de estos aspectos. En capítulos posteriores se aprenderá a calcular el número de repeticiones necesario para un experimento.

Al diseñar un experimento se debe tener en claro:

- que el experimento necesita repeticiones de las unidades experimentales.
- el tipo de experimento a realizar (manipulativo o mensurativo).
- que las unidades experimentales deben ser independientes.
- cuántas repeticiones debe tener el experimento.

A veces se registra la variable respuesta en parte de la unidad experimental, ya que demandaría gran esfuerzo (a veces innecesario) registrarla en toda la

unidad. A la porción de la unidad experimental en la que se mide la respuesta se la denomina unidad observacional. Es muy importante no confundir a la unidad observacional con la unidad experimental. La unidad experimental es aquella a la que se le aplica el tratamiento y constituye una repetición verdadera para ese tratamiento. Los experimentos deben contar con repeticiones verdaderas: unidades experimentales a las que se aplica repetidamente y de manera independiente el mismo tratamiento. En cambio el experimento puede correctamente no contener diferentes unidades observacionales en cada unidad experimental ni se requiere obligatoriamente definir la unidad observacional en experimentos en los cuales se registran las variables en toda la unidad experimental, ya que en muchas situaciones coinciden la unidad experimental y la unidad observacional.

1.9. Variabilidad: su control

Cuanto mayor sea la variación observada entre unidades que no resulte explicada por los tratamientos, por ejemplo la variación entre unidades que reciben el mismo tratamiento, menor será la capacidad del ensayo de detectar diferencias entre tratamientos (la potencia del ensayo). Dicho de otro modo, a mayor error experimental, mayor será el número de repeticiones necesario para alcanzar una determinada potencia. El rubro “error experimental” abarca todas las causas de variación que no se controlaron durante el ensayo. Entre éstas algunas no se controlaron porque son incontrolables por naturaleza. Todo material biológico, por homogéneo que sea, presenta una cierta fluctuación cuyos factores determinantes no se conocen y son por lo tanto incontrolables. Además, las observaciones están sujetas a errores de medición, que deben ser minimizados dentro de lo posible. En cambio, otras fuentes de variabilidad podrían ser controladas. Por ejemplo, en experimentos con animales se conoce que ciertos caracteres cuantitativos difieren entre el macho y la hembra, o entre animales de distintas camadas o de distinta edad o de diferente raza. En un diseño completamente al azar que utiliza animales de distintas edades, camadas, razas y sexos, toda esa variación contribuye al rubro del error.

Una alternativa para reducir el error experimental es aplicar los tratamientos sobre animales *tan homogéneos como sea posible* (lo mismo se aplica a parcelas, plantas, empresas o cualquier otro tipo de unidades experimentales). En esos casos se elimina la heterogeneidad *antes*, y se aplica un diseño completamente aleatorio. Sin embargo, se debe tener en cuenta que las conclusiones de tales ensayos sólo serán aplicables a condiciones tan restringidas como las que

se establecieron al seleccionar el material experimental. Es decir que *al acotar la variación del material que participa en el ensayo se restringe la población de referencia a la que apuntan las hipótesis*. Por ejemplo, se puede concluir que una ración produce el mayor aumento de peso en cerdos de determinada raza y edad.

Entonces, existe un compromiso entre:

- incluir un amplio rango de condiciones, con el fin de hacer el experimento más general.
- restringir el rango de condiciones, con el fin de reducir la variabilidad y aumentar la potencia.

Uno de los descubrimientos más originales de la estadística es que se puede eliminar del error experimental un factor controlable de variación *después* de haber realizado el ensayo usando material no homogéneo. Para que esto sea posible se debe conocer *a priori* la heterogeneidad entre unidades y diseñar consecuentemente un experimento con aleatorización restringida, por ejemplo con un **diseño en bloques**. Estos diseños permiten poner a prueba hipótesis referidas a poblaciones más heterogéneas y, en consecuencia, más amplias, sin provocar aumentos del error experimental que resten potencia al experimento. El más sencillo y más ampliamente utilizado de los diseños que controlan la heterogeneidad de esta manera se denomina **diseño en bloques completos** y se verá en detalle en los capítulos siguientes. Aquí sólo presentaremos su justificación conceptual.

Muchas veces el experimentador adquiere habilidad para predecir con cierta precisión el comportamiento de su material experimental: conoce que algunos lotes son más rendidores que otros, que algunos animales podrían aumentar de peso más rápidamente que otros, que alguna ubicación dentro del invernáculo recibe luz natural en un período más largo durante el día, etc. Todo conocimiento previo acerca de la heterogeneidad del material experimental puede servir para aumentar la potencia del experimento. En estos casos, el procedimiento a seguir comienza por clasificar a las unidades experimentales en conjuntos homogéneos y restringir la aleatorización de los tratamientos, aleatorizando por separado dentro de cada uno de esos grupos la asignación de todos los tratamientos. Cada grupo conformado por un conjunto de unidades experimentales homogéneas entre sí se denomina **bloque** y contiene una repetición completa de todos los tratamientos. Cuando se agrupan las unidades en bloques de manera tal que la variación entre las unidades de un bloque sea menor que la variación entre unidades de bloques diferentes, y se asignan al azar los tratamientos dentro de cada

bloque, es posible comparar las respuestas a los tratamientos entre unidades del mismo bloque. Así, las variaciones de un bloque al otro no contribuyen al error experimental, en consecuencia, la potencia del experimento aumenta como resultado de reducir el error experimental.

Desde un punto de vista matemático el tratamiento que se hace de los factores-bloque es el mismo que el de los factores-tratamiento, pero su significado dentro del modelo es diferente. Hay interés en conocer la influencia de un factor-tratamiento sobre la variable respuesta mientras que no hay interés en conocer la influencia de un factor-bloque sino que éste es incorporado al diseño del experimento para disminuir el efecto residual del modelo (el error experimental) y así aumentar la potencia.

Si la causa de la heterogeneidad es una propiedad cuantitativa de cada una de las unidades experimentales que puede ser medida antes de realizar el experimento (i.e. el peso inicial de los animales, la hora de la medición cuando se evalúa el estado hídrico de las plantas, la presión sanguínea de un paciente en un experimento médico, la acidez del suelo de las parcelas en un experimento de fertilización fosforada, etc.), el factor se denomina **covariable** y permite también disminuir el error experimental y aumentar la potencia. La manera de estimar y aislar el efecto de **covariables** la estudiaremos en el capítulo 7. Las posibles fuentes de variación de un experimento se pueden clasificar así:

FUENTE DE VARIACIÓN	TIPO
Debida a las condiciones de interés (Tratamientos)	Planificada y sistemática
Debida al resto de condiciones controladas (Bloques y covariables)	Planificada y sistemática
Debida a condiciones no controladas (Error experimental)	No planificada y aleatoria

1.10. Modelos estadísticos

Una vez establecido el objetivo del estudio, se elige un **modelo** que representa el comportamiento de la variable respuesta como función de las variables explicativas. El valor esperado de la respuesta se describe mediante una función matemática que relaciona a las variables explicativas con los parámetros poblacionales desconocidos. El modelo permite contrastar las predicciones del investigador con las observaciones del experimento (Fig. 1.2).

$$Respuesta\ observada = \underbrace{\text{modelo matemático}}_{f(\text{variables explicativas})} + \text{error experimental} \tag{1.1}$$

El modelo es la formulación de la distribución de probabilidad de una variable aleatoria, la variable respuesta, de manera que: 1) La esperanza de la variable respuesta es función matemática de las variables explicativas (los tratamientos o factores de interés), 2) los desvíos respecto a esas esperanzas son los errores experimentales independientes, que tienen distribución de probabilidad única, la que debe ser especificada como parte de la estructura del modelo. En este libro sólo se presentarán modelos que tienen una sola variable respuesta y de naturaleza cuantitativa. Primero veremos aquéllos donde la porción explicativa del modelo está compuesta por variables cualitativas, que también se conocen como tratamientos o variables de clasificación porque dividen a las observaciones en clases. A continuación se presentan algunos ejemplos.

Variable explicativa CUALITATIVA		Variable respuesta CUANTITATIVA
Cultivo antecesor	➔	Nivel de nutrientes en el suelo
Híbridos	➔	Rendimiento
Tipo de labranza	➔	Abundancia de malezas
Razas de terneros	➔	Producción anual de carne
Tipo de ordeño (manual/mecánico)	➔	Incidencia de mastitis en cabras
Calidad de la luz	➔	Germinación semillas de chamico
Especie forrajera	➔	Biomasa de la pastura
Nivel de educación	➔	Gasto en alimentos

Los modelos con variables predictoras cuantitativas, que se denominan *Modelos de regresión*, se estudiarán en los capítulos 6 y 7. A continuación se presentan algunos ejemplos.

Variable explicativa CUANTITATIVA		Variable respuesta CUANTITATIVA
Altura de los estudiantes	➔	Peso de los estudiantes
Precipitación media anual	➔	Productividad de pastizales
Cobertura del cultivo en floración	➔	Rendimiento del cultivo
Densidad de pastos en la estepa	➔	Producción anual de carne
Nivel de fertilidad del suelo	➔	Rendimiento de trigo
Intensidad de luz	➔	Tasa de fotosíntesis de la hoja
Tiempo desde emergencia	➔	Biomasa del cultivo
Diámetro del tronco	➔	Volumen de madera del árbol
Ingreso mensual familiar	➔	Gasto en alimentos

El modelo matemático debe indicar la relación que se supone que existe entre la variable respuesta y las principales fuentes de variación a estudiar. El modelo más habitual es el **modelo lineal**, como los siguientes:

Respuesta observada = Constante + Efecto de tratamientos + error (1.2)
modelo matemático

Si las unidades experimentales se agrupan en *bloques*:

Respuesta obs. = Constante + Efecto de bloques + Efecto de tratamientos + error (1.3)
modelo matemático

A veces los tratamientos que se comparan en un ensayo son, a su vez, combinaciones de niveles de distintos factores. Por ejemplo, si se desea estudiar el efecto de la fertilización nitrogenada sobre el nivel de proteína en grano en cultivos de trigo con y sin riego. En este caso, la fertilización es un factor que podría, por ejemplo, tener tres niveles (sin agregado de fertilizante, con dosis normal y con dosis alta) y el riego es otro factor que podría tener, por ejemplo, dos niveles (con y sin riego). Entonces las seis combinaciones de los niveles de ambos factores constituyen los tratamientos (Cuadro 1.1) y el porcentaje de proteína en grano es la variable respuesta.

Cuadro 1.1: En un experimento factorial los tratamientos son las combinaciones de los niveles de los factores incluidos en el ensayo

FACTOR 1: Fertilizante 3 Niveles	FACTOR 2: Riego 2 Niveles	TRATAMIENTOS (Combinaciones de los factores)
Sin agregado	Con riego	1. sin fertilizante- con riego
	Sin riego	2. sin fertilizante- sin riego
Dosis normal	Con riego	3. fertilización normal- con riego
	Sin riego	4. fertilización normal- sin riego
Dosis alta	Con riego	5. fertilización alta- con riego
	Sin riego	6. fertilización alta- sin riego

Los estudios de este tipo, donde la respuesta de un factor se prueba a diferentes niveles de otro factor en el mismo experimento se conocen como **experimentos factoriales**, son muy utilizados en las ciencias agropecuarias y presentan interesantes propiedades que estudiaremos más adelante. El modelo que representa la respuesta en un experimento factorial en el cual los tratamientos son *combinaciones de niveles de dos factores A y B* es:

Respuesta = Constante + Efecto de A + Efecto de B + Interacción AxB + error (1.4)
modelo matemático

En todos estos modelos la respuesta viene dada por una suma de términos que representan las principales fuentes de variación planificadas más un término residual debido a las fuentes de variación aleatoria no planificada. Si la variación explicada por el modelo y el residuo que queda sin explicar, constituyen fuentes independientes, se puede separar información de error. Obviamente el modelo será mejor si deja relativamente poco sin explicar. La intuición sugiere que el “mejor” modelo será el que arroje el menor promedio de errores o *residuales* (diferencia entre valores observados y valores predichos por el modelo).

En la formulación de cada modelo deben aparecer los **parámetros** poblacionales relevantes para el objetivo que impulsa el experimento, ya que el procedimiento que permite contestar las preguntas de interés es la prueba de hipótesis referida a los parámetros del modelo. De esta forma el modelo estadístico vincula los objetivos del estudio con el diseño del experimento y el análisis de los datos (Figura 1.8.). La estimación de los parámetros que caracterizan a la población de referencia permite evaluar la evidencia que proveen los datos respecto a las predicciones del investigador.

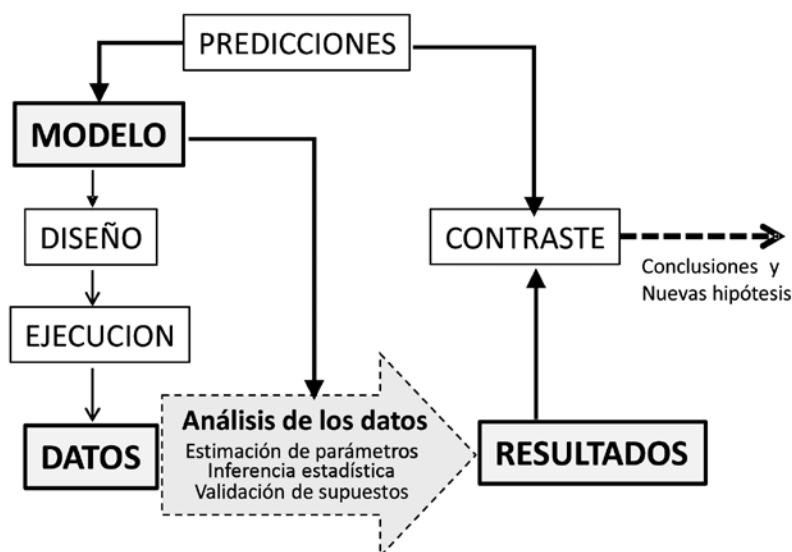


Figura 1.8: La complejidad del modelo y los parámetros a estimar son determinados por los objetivos del proyecto y al mismo tiempo son los determinantes de las fuentes de variación planificadas a ser consideradas en el diseño del experimento y en el análisis de los datos

Los parámetros poblacionales son elementos fundamentales del modelo estadístico que, como se verá en detalle más adelante, sintetizan de

manera formal las relaciones entre las variables bajo estudio. Es así que la selección del modelo estadístico más adecuado para cada caso estará determinada por los objetivos del estudio, y tanto el diseño como la conducción de los experimentos apuntarán a obtener la información necesaria para estimar correctamente los parámetros del modelo. Los modelos son “estadísticos” cuando tienen en cuenta la variación entre las unidades experimentales de la población y se pueden asociar probabilidades a los diferentes valores factibles de obtener. Estos conceptos serán discutidos en detalle en los próximos capítulos.

Asumiendo que se ha elegido el modelo apropiado para representar el comportamiento de la variable respuesta como función de las variables explicativas, entonces se pueden extraer ciertas conclusiones acerca de las hipótesis y extenderlas a la población de referencia, siempre y cuando la metodología del estudio incorpore la aleatorización y la repetición correspondientes al modelo y el análisis propuestos. Para que se pueda valorar la inseguridad de las conclusiones basadas en los datos mediante el cálculo de probabilidades, se deben cumplir los *supuestos* de cada método de análisis, que serán presentados en detalle en cada uno de los capítulos siguientes. Finalmente, el grado de cumplimiento de los objetivos del estudio dependerá de la definición y correspondencia entre la *población de referencia*, el *modelo*, el *muestreo* o *diseño*, el *análisis* y las *conclusiones*.

Caja 1.1: Medidas de variabilidad

Una manera frecuente de medir la variación de una variable es a través de los cuadrados de los desvíos con respecto al promedio. La suma de los cuadrados de estos desvíos (SC_y) tiene propiedades matemáticas interesantes, es una medida muy importante en estadística y a lo largo de este libro la invocaremos frecuentemente.

$$SC_y = \sum_{i=1}^n (y_i - \bar{y})^2$$

La VARIANZA poblacional (σ^2) es un valor que cuantifica la dispersión de la variable en estudio como el promedio de los cuadrados de los desvíos de la variable respecto de su media. El estimador insesgado de la varianza σ^2 es s^2 , a veces denominado varianza muestral.

$$S^2_y = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

La DESVIACION ESTANDAR (σ) es la raíz de la varianza, tiene las mismas unidades que la variable bajo estudio, su estimador es S_y . Los estimadores son variables aleatorias cuyos valores varían entre muestra y muestra, en consecuencia, también tienen varianza y desviación estándar.

La VARIANZA DE LA MEDIA MUESTRAL es el promedio de los cuadrados de los desvíos de las medias muestrales de todas las muestras de un cierto tamaño que pueden ser extraídas de una población, respecto de su promedio (que es igual a la media poblacional). La dispersión de las medias es considerablemente menor que la dispersión de la variable original, de hecho el valor de la varianza de la media muestral es el cociente entre la varianza de la variable en cuestión y el tamaño de la muestra.

$$Var(\bar{y}) = \frac{1}{n} Var(y)$$

El estimador de la desviación estándar de la media se denomina ERROR ESTANDAR de la media y se calcula con la información de una sola muestra de tamaño n

$$EE(\bar{y}) = \frac{s}{\sqrt{n}} = \sqrt{\frac{s^2}{n}}$$

Se invita al lector a repasar estos conceptos en un texto de estadística (Batista 2018)

1.11. Las pruebas de hipótesis estadísticas

Al poner a prueba una hipótesis estadística, el interés está centrado en someter a prueba una suposición de que la diferencia verdadera en las respuestas tiene algún valor específico, siendo cero el caso más común. El cero (que indica que no hay diferencias entre tratamientos) es la hipótesis nula que dice “nada interesante está ocurriendo”. Si los datos aportaran evidencias para refutarla estamos en condiciones de concluir que hay un efecto significativo, equivalente a decir que los tratamientos difieren en sus efectos. La dificultad surge porque, como resultado de su variabilidad, los datos nunca están exactamente de acuerdo con las hipótesis y, entonces, el problema reside en decidir si la discrepancia entre los datos y la hipótesis va a ser atribuida a esas variaciones o al hecho de que la hipótesis sea falsa.

Las reglas para decidir si una hipótesis se va a considerar falsa están construidas para satisfacer dos condiciones claramente deseables: (i) las hipótesis

que sean ciertas serán rechazadas sólo ocasionalmente y la probabilidad de que esto ocurra puede ser fijada previamente y, (ii) las hipótesis que sean falsas serán rechazadas tan frecuentemente como sea posible.

Una propiedad útil de este procedimiento es su influencia moderadora sobre aquel experimentador que tiende a extraer conclusiones fundamentadas en datos insuficientes y pretende convencer a los demás acerca de un efecto marcado de un tratamiento. El riesgo de rechazar erróneamente la hipótesis de *no efecto* cuando es correcta (probabilidad de error de tipo I), se denomina **nivel de significación** de la prueba. Las probabilidades más frecuentemente utilizadas para este riesgo son 0.05 y 0.01 pero estos niveles son únicamente convenciones útiles y pueden usarse probabilidades más bajas si fueran muy serias las consecuencias de un rechazo erróneo de la hipótesis.

El nivel de significación de una prueba pierde importancia cuando los datos de un experimento no aportan suficiente evidencia para rechazar la hipótesis nula. Por ejemplo, si el resultado de un estudio fuera “no se detectaron diferencias en la función hepática entre personas tratadas con la dosis más alta de un nuevo analgésico para calmar dolores de cabeza y aquellas que recibieron un placebo”, ¿eso significa que el analgésico es inocuo para el funcionamiento del hígado, aún en su dosis más alta?

¿Se puede suponer que la hipótesis nula es verdadera e interpretar que el efecto es cero? Evidentemente no. Cuando un investigador comunica que no encontró diferencias significativas, necesitamos que nos informe qué potencia tenía su prueba, es decir qué probabilidad de evitar un error de tipo II (aceptar una hipótesis nula que es en realidad falsa).

Caja 1.2: Algunas definiciones importantes

Hipótesis nula (H_0): Hipótesis que generalmente indica que no existen diferencias.

Hipótesis alternativa (H_1): Hipótesis que indica que existen diferencias.

Error Tipo I: Rechazo de una hipótesis nula verdadera.

Error Tipo II: Aceptación de una hipótesis nula falsa.

Alfa: Probabilidad de cometer un Error Tipo I.

Beta: Probabilidad de cometer un Error Tipo II.

Potencia: Capacidad de una prueba estadística para rechazar la hipótesis nula cuando no es verdadera. Probabilidad de no cometer error Tipo II.

1.12. Experimentos manipulativos y mensurativos

Alcances de las conclusiones: los análisis estadísticos para experimentos manipulativos y mensurativos generalmente son los mismos, la principal diferencia entre ambos surge al atribuir causas a las respuestas observadas (Hulbert 1984). En los experimentos manipulativos, uno puede asociar causas con tratamientos, es así que la variación en la variable respuesta se “particiona” en efecto de la variable independiente (tratamientos) y efecto de todos los otros factores o error experimental (Figura 1.5).

En cambio, los experimentos mensurativos no permiten separar fácilmente al factor bajo estudio de todas las otras causas (o fuentes) de variación. Por ejemplo, en un estudio de sociología rural, se deseaba conocer los efectos de la “Flexibilización laboral” aplicada durante la década del 90, sobre el número de trabajadores empleados que acceden a los beneficios (e.g., obra social, seguro de trabajo, jubilación, indemnización, aguinaldo). Para ello, se relevó el número de trabajadores con descuento jubilatorio en cinco departamentos de la provincia de Córdoba, cuya principal actividad productiva es el agro. Los relevamientos se realizaron en dos momentos (tratamientos): antes de la aplicación de la “Flexibilización laboral” (año 1990) y después de su aplicación (año 2003). Una vez obtenidos los datos, el análisis estadístico podrá mostrar si en el año 1990 hubo más o menos trabajadores con descuento jubilatorio que en el año 2003, y asociar una probabilidad de error a la conclusión elegida. Sin embargo, en caso de detectar una disminución en los trabajadores con descuento jubilatorio en el año 2003, esta diferencia no puede ser atribuida exclusivamente a la flexibilización laboral. Esto se debe a que otros factores (globalización, convertibilidad, etc.) que afectan a la variable respuesta también han cambiado durante este período. Es decir, no se puede separar el efecto de la aplicación de la flexibilización laboral de estos otros factores. La causalidad de la respuesta observada no puede ser atribuida exclusivamente al factor de interés (la flexibilización laboral). En cambio, si se puede separar el efecto de los tratamientos y los otros factores que “covarían” con ellos, de la variación observada entre los distintos partidos de la provincia, que son incluidos en el error experimental. Por lo tanto, la partición de la variabilidad en un experimento mensurativo es entre el efecto de los tratamientos junto con otros posibles factores que varíen conjuntamente con ellos y el error experimental (Figura 1.9.).

Es importante diseñar el experimento mensurativo de manera de poder separar la mayor cantidad de efectos de otros factores que puedan confundirse con el efecto de los tratamientos. Sin embargo, a diferencia de los experimentos manipulativos, nunca se podrá separar completamente el efecto de los factores de interés de todos los factores restantes. Esto no significa que los experimen-

tos mensurativos sean “malos” experimentos y los experimentos manipulativos sean “buenos”. Los experimentos manipulativos permiten atribuir la proporción de la variación total que se debe exclusivamente al factor bajo estudio, pero muchas veces carecen de realismo. En cambio, los experimentos mensurativos permiten estudiar los sistemas de interés directamente. Es por esto que se necesitan de los dos tipos de experimentos para avanzar en el conocimiento, y que generalmente se incluyen diferentes combinaciones de experimentos mensurativos y manipulativos en los programas científicos. Más aún, muchas veces no es ético realizar experimentos manipulativos (e.g. con personas) y gran parte de nuestro conocimiento se basa en experimentos mensurativos. Lo importante es que, ya sea que tengamos que basarnos solamente en experimentos mensurativos, no utilicemos análisis estadísticos incorrectamente para extraer conclusiones más allá de las que el tipo de experimento permite.

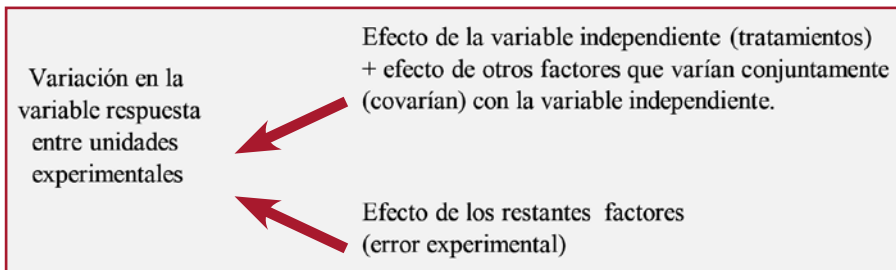


Figura 1.9: Determinantes de la variación entre distintas unidades en la variable respuesta en experimentos mensurativos cuando existen factores que covarían con la variable independiente

EJERCICIOS DE APLICACIÓN

Cuando hayas terminado con esta práctica podrás:

- Diseñar un experimento
- Conocer la importancia de las repeticiones para generar conclusiones válidas sobre los tratamientos
- Controlar las fuentes de error experimental que oscurecen los resultados
- Manejar la potencia estadística de un diseño experimental
- Construir modelos sencillos para explicar el comportamiento de una variable de interés.

1.1. Usted ha sido contratado por una empresa de agroquímicos para que estudie cuál de las tres formulaciones químicas de un fungicida (a las que denominaremos F1, F2 y F3), es más eficaz en el control de la “roya asiática” en cultivos de soja en el noreste argentino. El terreno experimental que tiene la empresa en la zona presenta

considerable homogeneidad ambiental (fertilidad del suelo, topografía, historia de uso, etc.).

- a) ¿Cuál es el objetivo del trabajo para el cual ha sido contratado?
- b) Plantee un ensayo para cumplir con dicho objetivo. Identifique el tipo de diseño experimental.
- c) ¿Cuáles son los tratamientos que usted probaría? (Tenga en cuenta que en la zona no se sabe si el fungicida puede ser realmente efectivo).
- d) ¿Qué variable/s respuesta registraría? Fundamente su elección.
- e) ¿En función de qué información decidiría el número de repeticiones del experimento?
- f) Dibuje un plano de campo con una posible asignación de los tratamientos (suponga cinco repeticiones).
- g) Evidentemente sería más sencillo dividir el lote en cuatro grandes franjas y aplicar un tratamiento en cada una de ellas. Argumente para convencer al director del campo experimental de la necesidad de conducir el experimento con repeticiones independientes. (Ayuda: ¿qué información se perdería con el ensayo más simple de grandes franjas?, ¿cuál es la necesidad de contar con esa información?)

1.2. A continuación se describen casos reales de estudios donde se abordan diferentes problemáticas. Para cada uno determine cuál es la variable independiente o explicativa y cuál es la variable dependiente o respuesta. Además, indique qué tipo de variable es cada una, cuáles son los tratamientos y cuáles las unidades experimentales. Esquematice en un gráfico una predicción de los resultados. Finalmente, explique si son experimentos manipulativos o mensurativos.

- a) Una productora hortícola de La Plata desea estudiar la relación entre el rendimiento (kg/planta) de lechuga capuchina (*Lactuca sativa*) y la dosis de lombricompost (0 - 0,5 - 1 - 1,5 - 2 kg/m²). Con este fin, realiza un ensayo con 50 camellones (5 plantas/camellón) y a cada uno le asigna aleatoriamente alguna de las 5 dosis de lombricompost.
- b) Un investigador desea estudiar la relación entre la disponibilidad de fósforo en el suelo y los años de siembra de monocultivo de soja en lotes agrícolas del norte de la provincia de Buenos Aires. Para realizar el estudio seleccionó 15 lotes al azar en los que se haya sembrado soja desde 1 hasta 9 años consecutivos. En cada lote midió el contenido de fósforo en el suelo. Para cada tiempo de monocultivo dispuso de datos de tres lotes independientes.
- c) Para estudiar si el tipo de producción (convencional u orgánica) influye sobre la diversificación productiva (número de especies de plantas y animales producidos) se entrevistaron 50 familias vinculadas a la producción convencional, 50 familias que producen de modo orgánico, y 50 familias que producen alternativamente en forma convencional y orgánica, seleccionadas aleatoriamente de diferentes departamentos de la provincia de Misiones.
- d) El lodo de desagüe es un residuo seco que se usa como fertilizante en la agricultura, siempre que no contenga niveles tóxicos de metales pesados. Un productor desea saber si los niveles de zinc (metal pesado) en sus cultivos

están relacionados con cantidades crecientes de lodo (tn/ha). La experiencia consistió en registrar niveles de zinc (en ppm) en el suelo en 21 parcelas con plantas de cebada que fueron fertilizadas con 3 cantidades distintas de lodo de desagüe provenientes del área metropolitana de Luján.

- 1.3. En una comunidad campesina de Santiago del Estero se desea estudiar el efecto de un balanceado proteico sobre la tasa de reproducción de las cabras (n° hijos/cabra x año). El balanceado proteico se formularía a partir de hojas y frutos de árboles del monte santiagueño. A cargo del ensayo se encuentran dos ingenieros agrónomos, uno especialista en nutrición animal y otro especialista en extensión rural. El nutricionista argumenta que para disminuir la variabilidad solo se deberían tomar animales de edades con diferencias menores al año (es decir entre 2 y 3 años, o entre 3 y 4 años, etc.). El extensionista no está de acuerdo, argumenta que que si toman rangos de edades como los propuestos, los resultados del ensayo solo serán aplicables a cabras de esos rangos de edades. El extensionista desea obtener información que abarque más condiciones en las conclusiones.
 - a) ¿Cuál es la unidad experimental y la unidad observacional bajo análisis?
 - b) Fundamente los argumentos del nutricionista y del extensionista utilizando conceptos del diseño de experimentos y la estadística inferencial.
 - c) Al parecer la discusión entre ambos sigue y no se pondrán de acuerdo, encuentre un diseño que pueda satisfacer a los dos ingenieros. Escriba con palabras el modelo que representa la partición de la variación observada en diferentes fuentes.
 - d) ¿En qué tipo de experimento están pensando ambos ingenieros: manipulativo o mensurativo? Justifique su respuesta
- 1.4. El Dr. Green se propone estudiar el efecto de la aplicación de un fertilizante sobre el rendimiento de un cultivo. Diseña un experimento con dos tratamientos (con y sin aplicación del fertilizante) y cinco repeticiones. Indica a su ayudante de campo cuáles son las parcelas que deben ser fertilizadas, según el resultado de la aleatorización. El ayudante tiene un percance que le impide hacer la aplicación del fertilizante, pero el Dr. Green nunca se entera que el tratamiento no fue aplicado. Después de la cosecha se comparan los rendimientos. La hipótesis de igualdad de medias entre parcelas “tratadas” y no tratadas se rechaza con un nivel de significación del 5%. Se concluye que el fertilizante produce un incremento significativo del rendimiento del cultivo. Explique este resultado.
- 1.5. Mencione cómo desde el punto de vista del diseño de un ensayo experimental puede:
 - reducir las fuentes del error experimental.
 - modelar las variables que contribuyen a aumentar la variabilidad en los resultados pero que no son interés de la investigación.
 - obtener una mejor estimación del error experimental.
 - Abarcar una población de referencia mayor