GRUPO BLR

DISEÑO EXPERIMENTAL EN MEDICINA VETERINARIA

PRINCIPIOS, DISEÑOS CLÁSICOS Y APLICACIONES EN INVESTIGACIÓN DE ESPECIES

ISBN: 978-9907-852-10-3

Luiggi Óscar Solano Maza
Ronald Andreé Vitonera Rogel

2026

DISEÑO EXPERIMENTAL EN MEDICINA VETERINARIA PRINCIPIOS, DISEÑOS CLÁSICOS Y APLICACIONES EN INVESTIGACIÓN DE ESPECIES

Luiggi Óscar Solano Maza

Ronald Andreé Vitonera Rogel

AUTORES



UTMACH

Este libro ha sido debidamente examinado y valorado en la modalidad doble par ciego con fin de garantizar la calidad científica.

©Grupo Editorial BLR
Universidad Técnica de Machala
Riobamba – Ecuador
Correo: publicaciones@grupoblr.com
https://grupoblr.com/libros-investigo
REPOSITORIO



Solano, L., Vitonera., R. (2026) Diseño experimental en medicina veterinaria Principios, diseños clásicos y aplicaciones en investigación de especies. Grupo Editorial BLR.

© Luiggi Óscar Solano Maza
Ronald Andreé Vitonera Rogel

Diseño experimental en medicina veterinaria Principios, diseños clásicos y aplicaciones en investigación de especies.

ISBN: 978-9907-852-10-3

El copyright promueve la libertad de expresión, protege la diversidad de ideas y conocimiento, además apoya la libre expresión. Se prohíbe de manera rigurosa la producción o el almacenamiento de esta publicación, ya sea en su totalidad o en parte, está estrictamente prohibido por ley, incluyendo el diseño de la portada, así como su difusión a través de cualquiera de sus medios, ya sean electrónicos, mecánicos, ópticos, de grabación o incluso de fotocopia, sin permiso de los propietarios de los derechos de autor.

Filiación de autores

Luiggi Óscar Solano Maza

Universidad Técnica de Machala

Correo Electrónico: losolano@utmachala.edu.ec

ORCID: https://orcid.org/0000-0002-8629-7516

Ronald Andreé Vitonera Rogel

Universidad Técnica de Machala

Correo Electrónico: rvitonera1@utmachala.edu.ec

ORCID: https://orcid.org/0000-0002-5272-5551

PREFACIO

Diseño Experimental en Medicina Veterinaria nace de una necesidad recurrente en la formación profesional, la de aprender a pasar de la observación clínica y productiva a la evidencia confiable, dado que en el campo veterinario decidir entre tratamientos, dietas, protocolos de manejo o estrategias de control sanitario exige más que intuición: reclama datos obtenidos con rigor, analizados con método e interpretados con criterio biológico, de tal manera que esta obra ha sido concebida para acompañar ese tránsito, integrando la lógica estadística del diseño experimental con problemas reales de la práctica veterinaria.

El propósito central del libro es ofrecer a estudiantes y profesionales una guía estructurada para formular preguntas investigables, diseñar experimentos válidos y extraer conclusiones útiles para la toma de decisiones, y a lo largo del texto se enfatiza que un resultado estadísticamente significativo solo adquiere valor cuando está respaldado por una unidad experimental bien definida, un plan de aleatorización coherente, un número de réplicas suficiente y una interpretación contextualizada en la biología de la especie, línea en la cual la obra promueve una formación que articula precisión cuantitativa, razonamiento crítico y responsabilidad científica.

La organización del con*t*enido responde a una secuencia didáctica progresiva, en la cual el Capítulo 1 establece los fundamentos del diseño experimental, con énfasis en réplica, aleatorización, control local, definición de unidades y prevención de pseudorreplicación,

mientras el Capítulo 2 desarrolla la inferencia estadística aplicada al experimento: prueba de hipótesis, errores tipo I y II, prueba t, ANOVA y comparaciones múltiples. A los diseños clásicos de un factor (DCA, DBCA y Cuadrado Latino) se dedica el Capítulo 3, en tanto que el Capítulo 4 amplía el horizonte hacia diseños factoriales, interacciones y diseños con restricciones de aleatorización, como parcela dividida, incorporando criterios prácticos para elegir el diseño más adecuado.

Desde el punto de vista pedagógico, cada capítulo integra objetivos de aprendizaje, preguntas de enfoque, desarrollo conceptual, ejemplos resueltos con aritmética explícita, aplicaciones veterinarias, actividades de consolidación y glosario, arquitectura con la que se busca que el lector no solo recuerde fórmulas, sino que comprenda cuándo, por qué y bajo qué supuestos debe aplicarlas, y las tablas, flujogramas y diagramas se han incluido como herramientas de decisión, no como adornos, con el fin de facilitar la transferencia de lo aprendido a situaciones concretas de clínica, producción e investigación.

El libro asume, además, una postura metodológica clara: diseñar bien es una exigencia ética y científica, ya que un experimento mal planteado consume recursos, tiempo y animales sin generar conocimiento robusto, razón por la cual se insiste en la verificación de supuestos, el control del error experimental, la transparencia del reporte y la distinción entre significancia estadística y relevancia biológica; en síntesis, esta obra pretende contribuir a la formación de médicos veterinarios capaces de evaluar críticamente la evidencia, producir

investigación reproducible y sustentar decisiones con base en datos sólidos.

Esperamos que este texto sea útil tanto en el aula como en el trabajo de campo y laboratorio. Si al finalizar su lectura el lector logra formular mejores preguntas, diseñar experimentos más rigurosos e interpretar resultados con mayor criterio, el propósito de esta obra se habrá cumplido.

ÍNDICE

PREFACIO ..... **i**

ÍNDICE ..... **iv**

ÍNDICE DE TABLAS ..... **xiii**

ÍNDICE DE FIGURAS ..... **xv**

CAPÍTULO I **20**

1 FUNDAMENTOS DEL DISEÑO EXPERIMENTAL **20**

1.1 Introducción 20

1.2 Objetivos de aprendizaje 21

1.3 Preguntas de enfoque 22

1.4 Generalidades del diseño experimental 23

1.4.1 ¿Qué es un experimento? 23

1.4.2 Historia breve del diseño experimental 26

1.4.3 ¿Por qué diseñar experimentos? 29

1.4.4 El método científico y el diseño experimental 31

1.4.5 Etapas del diseño experimental 33

1.5 Principios fundamentales del diseño experimental 35

1.5.1 Réplica……..36

1.5.2 Aleatorización..38

1.5.3 Control local (bloqueo) ..40

1.5.4 Interrelación entre los tres principios..44

1.6 Requisitos básicos de la experimentación..44

1.6.1 La unidad experimental..44

1.6.2 Unidad experimental vs. unidad de observación: pseudorreplicación..45

1.6.3 Tratamiento y factor..50

1.6.4 Variable respuesta..51

1.6.5 Error experimental..52

1.6.6 Hipótesis nula y alternativa..53

1.7 Las variables en la investigación científica..55

1.7.1 Clasificación de variables ..55

1.7.2 Escalas de medición..57

1.8 Medidas de resumen de datos ..58

1.8.1 Medidas de tendencia central..59

1.8.2 Medidas de dispersión..60

1.8.3 Medidas de posición 64

1.8.4 Representación gráfica 64

1.8.5 Presentación estándar de resultados 66

1.9 Conceptos clave 68

1.10 Actividades de consolidación 69

1.11 Ejercicios propuestos 70

1.12 Preguntas de autoevaluación 73

1.13 Glosario…… 77

1.14 Resumen…. 82

CAPÍTULO II 84

2 PRUEBA DE HIPÓTESIS Y ANÁLISIS DE VARIANZA . 84

2.1 Introducción 84

2.2 Objetivos de aprendizaje 85

2.3 Preguntas de enfoque 86

2.4 Prueba de hipótesis 86

2.4.1 Concepto de hipótesis estadística 87

2.4.2 Errores tipo I y tipo II 88

2.4.3 Nivel de significancia y valor p 91

2.4.4 Potencia estadística 92

2.4.5 Procedimiento general de la prueba de hipótesis 93

2.4.6 Prueba t de Student..... 94

2.4.7 Prueba t para dos muestras independientes 95

2.4.8 Prueba t para muestras pareadas 98

2.5 Análisis de varianza (ANOVA) 100

2.5.1 ¿Por qué ANOVA y no múltiples pruebas t? 101

2.5.2 Lógica del ANOVA: descomposición de la variabilidad..... 102

2.5.3 Fórmulas del ANOVA para un factor 105

2.5.4 Tabla ANOVA 107

2.5.5 Distribución F y región de rechazo 107

2.5.6 Supuestos del ANOVA 108

2.5.7 Comparaciones múltiples (post-hoc) 110

2.5.8 Reporte de resultados del ANOVA 112

2.6 Ejemplo integrador: antiparasitarios en ovinos 113

2.6.1 Caso clínico 113

2.7 Tratamiento antiparasitario 116

2.7.1 Segundo ejemplo: promotores de crecimiento en pollos 117

2.8 Flujograma de decisión estadística..... 120

2.9 Conceptos clave 120

2.10 Actividades de consolidación..... 122

2.11 Ejercicios propuestos 123

2.12 Preguntas de autoevaluación 125

2.13 Glosario…… 130

2.14 Resumen…… 134

CAPÍTULO III 136

3 DISEÑOS EXPERIMENTALES CLÁSICOS 136

3.1 Introducción 136

3.2 Objetivos de aprendizaje 137

3.3 Preguntas de enfoque 138

3.4 Diseño completamente al azar (DCA) 138

3.4.1 Definición y características 139

3.4.2 Modelo estadístico lineal 141

3.4.3 Aleatorización en el DCA 142

3.4.4 Fórmulas del ANOVA para el DCA 143

3.4.5 Ejemplo: evaluación de 4 antiparasitarios en caprinos 145

3.4.6 Ventajas y limitaciones del DCA 147

3.5 Diseño en bloques completos al azar (DBCA) 148

3.5.1 ¿Por qué bloquear? 148

3.5.2 Modelo estadístico lineal 149

3.5.3 Aleatorización en el DBCA 150

3.5.4 Tabla ANOVA y fórmulas del DBCA 151

3.5.5 Ejemplo: suplementos minerales en vacas lecheras 153

3.6 Diseño en cuadrado latino (DCL) 155

3.6.1 ¿Cuándo usar un cuadrado latino? 155

3.6.2 Modelo estadístico lineal 156

3.6.3 Estructura del cuadrado latino 157

3.6.4 Tabla ANOVA para el DCL 158

3.6.5 Ejemplo: tratamientos para mastitis bovina 159

3.7 Comparación de los tres diseños 160

3.8 Conceptos clave 163

3.9 Actividades de consolidación..... 164

3.10 Ejercicios propuestos 165

3.11 Preguntas de autoevaluación..... 166

3.12 Glosario…… 169

3.13 Resumen……...172

CAPÍTULO IV..... 173

4 EXPERIMENTOS FACTORIALES 173

4.1 Introducción 173

4.2 Objetivos de aprendizaje..... 174

4.3 Preguntas de enfoque 175

4.4 Experimentos factoriales: conceptos básicos..... 176

4.4.1 ¿Qué es un experimento factorial?..... 176

4.4.2 Terminología factorial..... 177

4.4.3 Efecto principal e interacción..... 179

4.4.4 Ventajas del diseño factorial 181

4.5 Modelo y ANOVA de la factorial a × b en DCA..... 183

4.5.1 Modelo estadístico lineal 183

4.5.2 Tabla ANOVA de la factorial 184

4.5.3 Fórmulas de sumas de cuadrados 185

4.5.4 Ejemplo: factorial 2 × 3 en nutrición porcina 186

4.5.5 Estructura del diseño 190

4.5.6 Modelo estadístico 191

4.5.7 Tabla ANOVA de la parcela dividida 192

4.5.8 Ejemplo: alojamiento × suplementación en vacas lecheras 194

4.6 Diseños especiales: visión panorámica 196

4.6.1 Medidas repetidas 196

4.6.2 Diseño cruzado (crossover) 197

4.6.3 Diseños anidados vs. cruzados 197

4.6.4 Diseños aumentados 198

4.7 Guía de selección de diseño 198

4.8 Conceptos clave 201

4.9 Actividades de consolidación 202

4.10 Ejercicios propuestos 203

4.11 Preguntas de autoevaluación 205

4.12 Glosario……..... 209

4.13 Resumen…… 212

REFERENCIAS BIBLIOGRÁFICAS 214

ÍNDICE DE TABLAS

Tabla 1. Identificación de la unidad experimental en diferentes contextos veterinarios..... 45

Tabla 2. Tipos de factores y ejemplos en investigación veterinaria..... 51

Tabla 3. Ejemplos de variables respuesta en investigación veterinaria..... 52

Tabla 4. Escalas de medición con ejemplos veterinarios y pruebas estadísticas apropia-das. 57

Tabla 5. Comparación de las medidas de tendencia central: cuándo usar cada una..... 59

Tabla 6. Formato estándar de presentación de resultados descriptivos (ejemplo: ganancia diaria de peso en pollos)..... 67

Tabla 7. Tipos de error en la prueba de hipótesis y sus consecuencias..... 89

Tabla 8. Inflación del error tipo I con múltiples pruebas t (α = 0,05 por prueba). 101

Tabla 9. Tabla Anova. 107

Tabla 10. Comparación de pruebas de comparaciones múltiples. 111

Tabla 11. Datos transformados. 114

Tabla 12. Ventajas y limitaciones del DCA. 147

Tabla 13. Comparación de los tres diseños. 161

Tabla 14. Número de combinaciones de tratamientos en diferentes factoriales. 178

Tabla 15. Estructura de la tabla ANOVA para un factorial a × b en DC. 184

Tabla 16. Comparación entre diseños anidados y cruzados. 198

Tabla 17. Resumen comparativo de todos los diseños experimentales del libro. 199

Tabla 18. Los datos de producción de leche (L/día). 204

ÍNDICE DE FIGURAS

Figura 1.Hitos fundamentales en la historia del diseño experimental. Desde la prueba t de Student (1908) hasta las guías ARRIVE (Kilkenny et al., 2010), el campo ha evolucionado hacia mayor rigor, reproducibilidad y consideración ética del uso de animales. 28

Figura 2. Etapas del método científico aplicado a la investigación veterinaria. La flecha punteada indica el ciclo de retroalimentación cuando los resultados requieren nuevas investigaciones. 32

Figura 3. Interrelación entre los tres principios fundamentales del diseño experimental. Cada principio refuerza la función de los otros dos. 43

Figura 4. Jerarquía entre unidad experimental (UE), unidad de observación (UO) y unidad de muestreo (UM). Si la dieta se aplica al corral, la UE es el corral, no el animal. 46

Figura 5. Comparación entre una asignación correcta (izquierda) y pseudorreplicación (derecha) en un ensayo de dietas en aves. A la izquierda, cada corral es una UE independiente; a la derecha, los pollos dentro de un solo corral no son independientes. 49

Figura 6. Clasificación de variables según su función en el experimento. Las flechas indican la dirección del efecto y la relación con la variable dependiente. 56

Figura 7. Ejemplo Resuelto..... 63

Figura 8. Ejemplo Resuelto 2..... 65

Figura 9.Diagrama de cajas (boxplot) comparativo de tres dietas en pollos de engorde. La caja representa el rango intercuartílico (Q1 a Q3), la línea horizontal dentro de la caja es la mediana, y los bigotes se extienden hasta los valores mínimo y máximo. 66

Figura 10.Relación entre errores tipo I (α), tipo II (β) y potencia (1 β). La zona sombreada roja es la región de rechazo bajo H0. La zona verde clara es β (probabilidad de no detectar un efecto real). La zona verde oscura es la potencia. 90

Figura 11. Procedimiento de cinco pasos para la prueba de hipótesis. El orden es estricto: los pasos 1 y 2 se definen antes de recolectar datos; los pasos 3-5 se ejecutan después. 94

Figura 12. Ejemplo resuelto 3..... 98

Figura 13. Ejemplo resuelto 4..... 100

Figura 14.Descomposición de la variabilidad total en el ANOVA. La prueba F evalúa si SCTrat es "suficientemente grande" en relación con SCError, usando los cuadrados medios como estimadores de varianza. 104

Figura 15. Distribución F con gl1 = 3 y gl2 = 16. Si Fcal > Ftab = 3,49, se rechaza H0 al nivel α = 0,05. La zona sombreada roja representa la probabilidad α de rechazar H0 cuando es verdadera. 108

Figura 16. Caso clínico parte 1. 115

Figura 17. Caso clínico parte 2. 116

Figura 18. Medias de OPG transformados por tratamiento antiparasitario (ordenadas de menor a mayor). Letras diferentes indican diferencias significativas (Tukey HSD, p < 0,05). Barras más bajas = mayor eficacia antiparasitaria 117

Figura 19. Ejemplo Resuelto parte 1. 118

Figura 20. Ejemplo Resuelto parte 2. 119

Figura 21. Flujograma de selección de prueba estadística para comparar medias de grupos independientes. Los supuestos se verifican con Shapiro-Wilk (normalidad) y Levene (homocedasticidad) (Kruskal & Wallis, 1952; Mann & Whitney, 1947). 120

Figura 22. Layout de un DCA con 3 tratamientos y 4 réplicas. Los tratamientos se distribuyen completamente al azar entre las 12 UE. Nótese la ausencia de agrupación o restricción.g 143

Figura 23. Estructura de la tabla ANOVA para el DCA. 144

Figura 24. Caso clinico parte 1. 145

Figura 25. Caso clínico parte 2. 146

Figura 26.Reducción porcentual de OPG por tratamiento antiparasitario en caprinos Boer (DCA). Letras diferentes indican diferencias significativas (Tukey HSD, p < 0,05). 147

Figura 27. Layout de un DBCA con 4 tratamientos y 5 bloques. Dentro de cada bloque, los tratamientos se asignan al azar. 151

Figura 28. Análisis de varianza para un diseño en bloques. 151

Figura 29. Caso Clínico. 153

Figura 30. Ejercicio. 154

Figura 31. Cuadrado Latino 4 4: 4 tratamientos de mastitis (A–D), filas = vacas, columnas = semanas. Cada tratamiento aparece exactamente una vez por fila y una vez por columna. 157

Figura 32. Tabla ANOVA para el DCL. 158

Figura 33. Caso clínico. 159

Figura 34. Ejemplo resuelto. 160

Figura 35. Flujograma de selección entre DCA, DBCA y Cuadrado Latino. 163

Figura 36. (a) Sin interacción 180

Figura 37. (b) Con interacción. 180

Figura 38. Caso Clínico. 186

Figura 39. Gráfico de interacción para la factorial proteína lisina. Las líneas aproximadamente paralelas confirman la ausencia de interacción significativa: el efecto de la lisina es consistente en ambos niveles de proteína. 187

Figura 40. Ejemplo Resuelto. 188

Figura 41. Estructura de un diseño en parcela dividida. El factor A (tipo de alojamiento) se aplica a parcelas principales; el factor B (suplementación) se aleatoriza dentro de cada parcela principal. Hay dos errores (a- entre parcelas) (b- dentro de sub-parcelas). 190

Figura 42. Tabla ANOVA. 192

Figura 43. Caso Clínico. 194

Figura 44. Ejercio Resuelto..... 195

Figura 45. Flujograma de selección de diseño experimental. La primera decisión es si se evalúa uno o más factores; la segunda, la estructura de variabilidad o las restricciones de aleatorización. 198

Figura 46. Aplicación Veterinaria. 200

CAPÍTULO I

1 FUNDAMENTOS DEL DISEÑO EXPERIMENTAL

1.1 Introducción

La investigación veterinaria enfrenta un desafío central, el de obtener conclusiones válidas a partir de datos que, por naturaleza, presentan variabilidad biológica, pues dos animales de la misma especie, raza y edad, sometidos al mismo tratamiento, rara vez responden de forma idéntica, y esa variabilidad inherente, producto de diferencias genéticas, fisiológicas, ambientales y de manejo, constituye el ruido de fondo contra el cual el investigador debe detectar señales reales de efecto. En este contexto, el **diseño experimental** es la disciplina que proporciona las herramientas para planificar investigaciones de modo que las diferencias observadas puedan atribuirse con confianza a los factores de interés y no al azar ni a variables confusoras (Montgomery, 2017).

Sin un diseño adecuado, ningún análisis estadístico posterior puede rescatar un experimento mal planificado, ya que, como señaló Fisher (1935) hace casi un siglo, la estadística no puede hacer decir a los datos lo que los datos no contienen, mientras que un diseño cuidadoso sí puede garantizar que los datos contengan la información necesaria para responder la pregunta de investigación, máxima que sigue siendo la piedra angular de toda investigación experimental moderna, desde ensayos clínicos farmacéuticos hasta pruebas de campo en producción animal (Bate & Clark, 2014;

Lazic, 2016).

En el contexto veterinario, el diseño experimental adquiere una dimensión ética adicional, dado que el uso de animales en investigación impone la obligación moral y legal de obtener la máxima información con el mínimo número de sujetos, siguiendo los principios de las 3Rs (Reemplazo, Reducción y Refinamiento) establecidos por Russell y Burch (1959) y reforzados por las guías ARRIVE (Percie Percie du Sert et al., 2020; Kilkenny et al., 2010), de tal manera que un diseño estadísticamente deficiente que requiere repetir el experimento con más animales viola el principio de Reducción, en tanto que uno con excesivas réplicas contraviene el principio de necesidad.

Este capítulo presenta los fundamentos conceptuales del diseño experimental: los principios de réplica, aleatorización y control local; la identificación correcta de la unidad experimental; la clasificación de variables; y las herramientas descriptivas básicas para resumir datos, conceptos que constituyen el cimiento sobre el cual se construyen los diseños experimentales formales que se estudiarán en los capítulos siguientes.

1.2 Objetivos de aprendizaje

1. Definir qué es un experimento y diferenciarlo de un estudio observacional.

2. Explicar los tres principios fundamentales del diseño experimental: réplica, aleato-rización y control local.

3. Identificar correctamente la unidad experimental, la unidad de observación y la unidad de muestreo en contextos veterinarios.

4. Clasificar variables según su función (independiente, dependiente, confusora) y su escala de medición.

5. Calcular e interpretar medidas de tendencia central, dispersión y posición con datos de investigación animal.

6. Formular correctamente hipótesis nula y alternativa para un problema de investi-gación veterinaria.

7. Reconocer la pseudorreplicación y explicar por qué invalida los resultados de un experimento.

8. Describir las etapas de planificación de un diseño experimental y su relación con el método científico.

1.3 Preguntas de enfoque

1. ¿Por qué es insuficiente observar que un tratamiento "parece funcionar" sin un diseño experimental formal?

2. ¿Cuál es la diferencia entre la unidad experimental y la unidad de observación, y por qué confundirlas invalida los resultados?

3. ¿Cómo protege la aleatorización contra sesgos que el investigador no puede anticipar?

4. ¿Qué información aporta el coeficiente de variación que la desviación estándar por sí sola no proporciona?

5. ¿Por qué es importante planificar el análisis estadístico *antes* de recolectar los datos?

6. ¿Cuántas réplicas son suficientes? ¿Existe un mínimo aceptable?

1.4 Generalidades del diseño experimental

1.4.1 ¿Qué es un experimento?

Un **experimento** es una investigación planificada en la que el investigador *manipula deliberadamente* uno o más factores (variables independientes), *controla* las condiciones experimentales y *mide* el efecto sobre una o más variables de respuesta, mientras mantiene constantes o aleatorizadas las demás fuentes de variación (Montgomery, 2017), definición formal en la cual se reconocen tres elementos esenciales que deben estar presentes simultáneamente para que una investigación califique como experimental.

El primer elemento es la **manipulación activa**: el investigador decide qué tratamientos aplicar, en qué dosis, durante cuánto tiempo y a quién, de manera que no se limita a observar lo que ocurre naturalmente, sino que interviene para crear las condiciones que le permitirán

responder su pregunta de investigación. El segundo elemento es el **control**: se mantienen constantes (o se registran y ajustan

estadísticamente) todas las variables que podrían afectar la respuesta pero que no son de interés, mientras que el tercero corresponde a la **asignación aleatoria**, en la cual los tratamientos se distribuyen entre las unidades experimentales mediante un mecanismo aleatorio que elimina sesgos de selección (Ott & Longnecker, 2016; Quinn & Keough, 2002).

La distinción respecto a un **estudio observacional** es fundamental, ya que en un estudio observacional el investigador registra lo que ocurre sin intervenir, mide, clasifica y analiza, pero no manipula; así, por ejemplo, registrar la prevalencia de mastitis en fincas lecheras según su sistema de ordeño es observacional, mientras que asignar aleatoriamente vacas a tres protocolos de sellado de pezón y medir la incidencia de mastitis es experimental, diferencia con implicaciones profundas para la inferencia causal, pues solo el experimento permite afirmar que una diferencia observada se *debe* al tratamiento (Dohoo et al., 2009; Thrusfield, 2018).

Concepto Clave

Un **experimento** requiere tres elementos: (1) manipulación deliberada de al menos un factor, (2) control de las condiciones experimentales, y (3) asignación aleatoria de las unidades experimentales a los tratamientos, de tal manera que, si falta alguno de ellos, no se trata de un experimento en sentido estricto (Fisher, 1935; Quinn & Keough, 2002).

En veterinaria, la frontera entre estudio observacional y experimental no siempre es nítida, y un estudio que compara la producción de leche entre fincas con ordeño manual vs. mecánico

parece experimental, aunque si esas fincas ya tenían su sistema de ordeño antes del estudio, no hubo asignación aleatoria: las que adoptaron ordeño mecánico pueden diferir sistemáticamente de las que no lo hicieron (tamaño, tecnificación, raza del ganado), lo cual introduce **confusión**, dado que no se puede separar el efecto del sistema de ordeño del efecto de las diferencias preexistentes entre fincas. El diseño experimental resuelve este problema porque la aleatorización garantiza que, en promedio, los grupos sean comparables en todas las variables, incluso en aquellas que el investigador no midió ni anticipó (Daniel & Cross, 2013; Petrie & Watson, 2013).

Aplicación

¿Experimento u observación? Identifique la naturaleza de cada estudio:

a) Se asignan aleatoriamente 30 terneros a 3 protocolos de calostro y se mide la concentración sérica de IgG → **Experimental** (manipulación + aleatorización).

b) Se comparan los perfiles hematológicos de perros con y sin leishmanio-sis, seleccionados de la casuística hospitalaria → **Observacional** (no hay manipulación ni aleatorización).

c) Se evalúan 4 antibióticos en placas de cultivo inoculadas con *E. coli* →

Experimental (manipulación + control + aleatorización).

d) Se registra la producción de huevos en granjas que voluntariamente adopta-ron iluminación LED vs. las que mantuvieron iluminación convencional →

e) **Observacional** (no hay asignación aleatoria).

1.4.2 Historia breve del diseño experimental

Los fundamentos del diseño experimental moderno tienen una historia fascinante que comienza en los campos agrícolas de Inglaterra y se extiende hasta los laboratorios de investigación veterinaria del siglo XXI.

Dato Curioso

En la década de 1920, **Ronald A. Fisher** (1890–1962), un joven estadístico contratado por la Estación Experimental de Rothamsted (Hertfordshire, Inglaterra), enfrentó un problema aparentemente simple: ¿qué variedad de trigo produce más en un campo determinado? La dificultad radicaba en que el campo no era homogéneo, pues el suelo variaba en fertilidad, humedad y textura de un metro a otro, mientras los rendimientos fluctuaban por causas desconocidas, ante lo cual Fisher comprendió que, para separar el efecto de la variedad del efecto de la heterogeneidad del campo, necesitaba un marco lógico y matemático riguroso, y así nació la tríada fundacional del diseño experimental: **aleatorización**, **bloqueo** y **análisis de varianza** (ANOVA). Su libro *The Design of Experiments* (1935) sigue siendo la referencia intelectual de nuestra disciplina (Fisher, 1935).

La obra de Fisher no surgió en un vacío, ya que antes de él los agrónomos de Rothamsted habían acumulado décadas de datos experimentales, aunque los analizaban con métodos que hoy consideraríamos ingenuos: promedios simples, sin estimación del error ni pruebas formales de significancia. Fisher introdujo tres innovaciones revolucionarias que transformaron la investigación empírica:

1. **La aleatorización como principio lógico** (no solo práctico): Fisher demostró que la asignación aleatoria de tratamientos es *necesaria* para que las pruebas de significancia tengan validez lógica. Sin aleatorización, el valor *p* carece de interpretación (Fisher, 1925).

2. **El análisis de varianza** (ANOVA): Fisher desarrolló el marco matemático que descompone la variabilidad total de un conjunto de datos en componentes atribuibles a distintas fuentes (tratamientos, bloques, error), de tal manera que resulta posible evaluar si las diferencias entre tratamientos superan lo esperado por azar.

3. **El diseño factorial**: Fisher fue el primero en proponer que se evaluaran *varios factores simultáneamente* en un solo experimento, en lugar de uno a la vez, lo que resulta más eficiente y permite detectar interacciones.

Tras Fisher, otros estadísticos ampliaron y sistematizaron el campo (Figura 1): Cochran y Cox (1957) publicaron *Experimental Designs*, el catálogo más completo de diseños experimentales de su época, que sigue siendo una referencia; el diseño experimental llegó al ámbito industrial de la mano de Box et al. (2005), con su enfoque en optimización y superficies de respuesta, mientras que Steel et al. (1997) adaptaron estos métodos al contexto de las ciencias biológicas y agropecuarias, con énfasis en aplicaciones de campo.

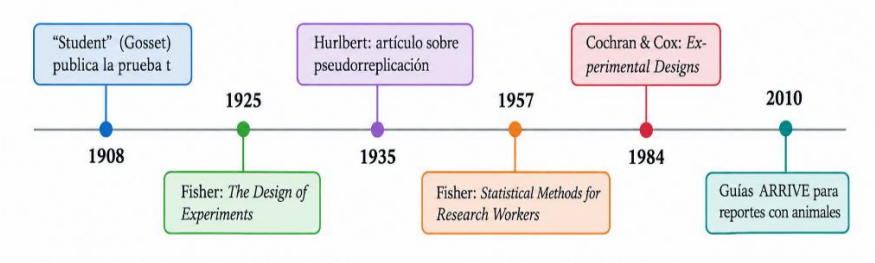


Figura 1.Hitos fundamentales en la historia del diseño experimental. Desde la prueba t de Student (1908) hasta las guías ARRIVE (Kilkenny et al., 2010), el campo ha evolucionado hacia mayor rigor, reproducibilidad y consideración ética del uso de animales.

En la investigación veterinaria, la adopción formal del diseño experimental fue más tardía que en agricultura, ya que durante décadas muchos estudios con animales se realizaban con diseños informales, muestras pequeñas y sin aleatorización explícita. El artículo seminal de Hurlbert (1984) sobre pseudorreplicación en ecología marcó un punto de inflexión, pues al documentar que el 48 % de los estudios publicados en revistas ecológicas de alto impacto cometían pseudorreplicación puso en evidencia la magnitud del problema, y ese mismo análisis lo extendió Lazic (2010) a la neurociencia, donde encontró errores similares; en veterinaria, Festing (2006) y Festing y Altman (2002) argumentaron que el diseño experimental deficiente no solo produce ciencia de baja calidad, sino que desperdicia vidas animales.

La preocupación actual por la reproducibilidad científica y el bienestar animal ha impulsado la formalización de estándares de reporte como las guías ARRIVE (Percie Percie du Sert et al., 2020; Kilkenny et al., 2010) y de herramientas de evaluación de sesgo como SYRCLE (Hooijmans et al., 2014), instrumentos que, a su

vez, exigen a los investigadores reportar explícitamente el diseño experimental, el método de aleatorización, el tamaño de muestra y su justificación, así como el análisis estadístico planificado. La era de los diseños informales ha terminado.

1.4.3 ¿Por qué diseñar experimentos?

Un experimento sin diseño formal es como una cirugía sin protocolo, ya que puede salir bien por suerte, pero los riesgos son inaceptables, de tal manera que el diseño experimental no es un lujo académico, sino una necesidad práctica con consecuencias reales para la validez de las conclusiones, el uso eficiente de recursos y el bienestar animal.

El diseño experimental permite:

a) **Maximizar la información** obtenida con un número dado de unidades experi-mentales. Un diseño eficiente extrae más información de los mismos datos que uno ineficiente (Casella, 2008; Mead et al., 2012).

b) **Minimizar el error experimental** mediante control de fuentes de variación conocidas. Técnicas como el bloqueo pueden reducir el error experimental a la mitad o menos, equivaliendo a duplicar el tamaño de muestra sin usar más animales.

c) **Evitar el sesgo** mediante aleatorización y enmascaramiento. El sesgo sistemá-tico no se reduce con más réplicas; solo la aleatorización lo neutraliza (Ott & Longnecker, 2016).

d) **Permitir inferencias causales** válidas, no solo asociaciones. La capacidad de atribuir un efecto a una causa es exclusiva del experimento controlado (Dohoo et al., 2009).

e) **Cumplir con principios éticos** (las 3Rs: Reemplazo, Reducción, Refinamiento), al usar el mínimo número de animales necesario para obtener resultados confiables (Festing et al., 2002; National Research Council, 2011; Russell & Burch, 1959).

f) **Facilitar la reproducibilidad**: un experimento bien diseñado y bien documenta-do puede ser replicado por otros investigadores, fortaleciendo la base de evidencia (Kilkenny et al., 2010).

Alerta Importante

Un error común en investigación veterinaria es analizar datos de un experimento mal diseñado y esperar que la estadística "arregle" los problemas. Ningún método estadístico puede compensar un diseño deficiente: si la aleatorización fue inadecua-da, si no hay réplicas suficientes, o si la unidad experimental está mal identificada, las conclusiones carecerán de validez sin importar cuán sofisticado sea el análisis (Hurlbert, 1984). Como dijo Fisher: "consultar al estadístico después de haber hecho el experimento es como pedirle que haga una autopsia: solo puede decirte de qué murió el experimento."

Considérese el costo de *no* diseñar correctamente: un investigador que no aleatoriza puede obtener un resultado aparentemente positivo que en realidad se debe a un sesgo de selección, porque los animales más sanos fueron asignados al tratamiento experimental, y publicarlo genera falsa evidencia, de tal manera que otros investigadores intentan replicar el hallazgo y fracasan, con el consiguiente desperdicio de recursos,

tiempo y animales, escenario que el diseño experimental previene desde el inicio (Pagano & Gauvreau, 2018; Wasserstein & Lazar, 2016).

1.4.4 El método científico y el diseño experimental

El diseño experimental opera dentro del marco del método científico. La investigación científica no es una colección de pasos mecánicos, sino un proceso lógico iterativo que busca explicaciones sobre fenómenos naturales mediante evidencia empírica controlada. La secuencia lógica general es:

- **Observación**: identificar un problema o fenómeno (ej.: alta mortalidad en terneros neonatos en una región).
- **Formulación de hipótesis**: proponer una explicación tentativa (ej.: la suple-mentación con calostro bovino hiperinmune reduce la mortalidad neonatal).
- **Diseño del experimento**: planificar cómo probar la hipótesis (selección de diseño, tamaño de muestra, variables a medir, método de aleatorización).
- **Ejecución**: aplicar los tratamientos y recolectar datos siguiendo el protocolo predefinido.
- **Análisis**: aplicar las pruebas estadísticas apropiadas, previamente planificadas.
- **Interpretación y conclusión**: decidir si la evidencia apoya o

rechaza la hipótesis, y contextualizar los hallazgos.

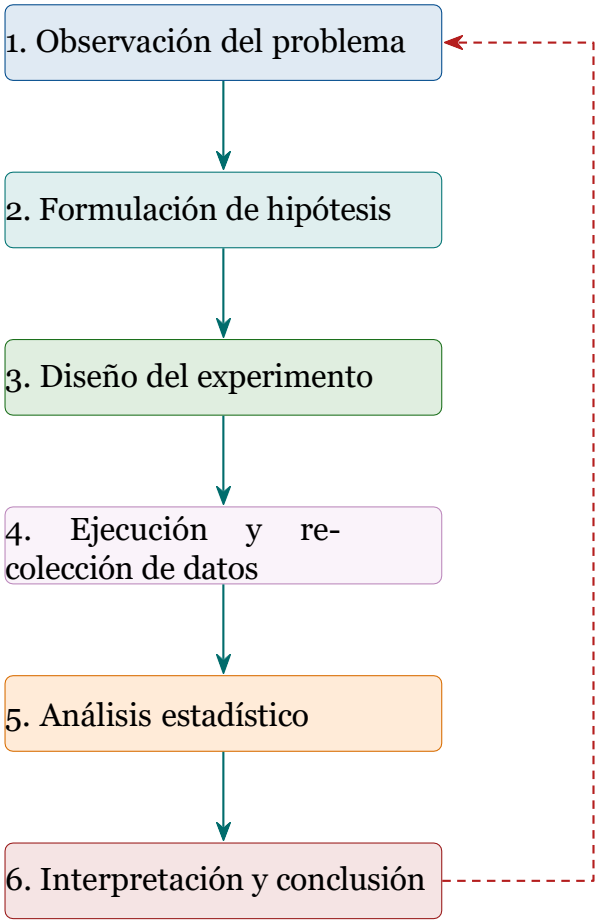


Figura 2. Etapas del método científico aplicado a la investigación veterinaria. La flecha punteada indica el ciclo de retroalimentación cuando los resultados requieren nuevas investigaciones.

La etapa 3 (diseño) no es un paso que se pueda saltar o improvisar, ya que todo lo que sigue, esto es, la ejecución, el análisis y la interpretación, depende de las decisiones tomadas en el diseño: un diseño incorrecto produce datos que ningún análisis puede rescatar,

mientras que uno eficiente arroja datos que se analizan fácilmente y responden la pregunta con claridad; en palabras de Montgomery (2017): “un análisis estadístico siempre puede hacerse; la pregunta es si tendrá significado.”

1.4.5 Etapas del diseño experimental

La etapa de diseño (paso 3 de la Figura 2 bookmark11) se descompone en decisiones secuenciales que Montgomery (2017) resume en siete pasos:

- Definir claramente el objetivo del experimento.
- Identificar los factores y sus niveles.
- Seleccionar la variable respuesta y su método de medición.
- Elegir el diseño experimental (DCA, DBCA, cuadrado latino, factorial, etc.).
- Determinar el número de réplicas.
- Planificar el orden de ejecución (aleatorización).
- Anticipar el análisis estadístico.

Este orden no es arbitrario: decidir el análisis *antes* de recolectar datos es una de las mejores prácticas para evitar el análisis exploratorio post-hoc no justificado. La tentación de “buscar significancia” probando múltiples análisis después de recolectar datos, lo que se conoce como *p-hacking* o “dragado de datos”, es un

problema grave en la ciencia contemporánea (Nakagawa & Cuthill, 2007; Wasserstein & Lazar, 2016). Planificar el análisis a priori protege contra esta práctica.

Concepto Clave

El diseño experimental debe planificarse **antes** de recolectar datos, ya que las siete decisiones del diseño (objetivo, factores, respuesta, tipo de diseño, réplicas, aleatorización y análisis) forman un todo coherente, en el que modificar una de ellas puede requerir ajustar las demás (Kuehl, 2000; Montgomery, 2017).

Ejemplo Resuelto

Planificación completa de un experimento veterinario.

Situación: Un médico veterinario desea evaluar si un nuevo probiótica mejora la ganancia de peso en pollos de engorde. ¿Cómo planifica el experimento?

Solución paso a paso:

1. **Objetivo:** Determinar si el probiótico X mejora la ganancia diaria de peso (GDP) en pollos de engorde Ross 308 durante la fase de crecimiento (1–21 días).

2. **Factores y niveles:** Un solo factor (suplementación) con 3 niveles: con-trol (sin probiótico), dosis baja (0.5 g/kg alimento), dosis alta (1.0 g/kg alimento).

3. **Variable respuesta:** GDP (g/día), calculada como (peso final peso inicial) / 21 días. Medición: balanza digital

con precisión de 1 g.

4. **Diseño:** DCA (Diseño Completamente al Azar), ya que los pollos provienen de la misma incubadora, misma edad y peso inicial similar (homogéneos).

5. **Réplicas:** 5 corrales por tratamiento (el corral es la UE si se mezcla el probiótico en el comedero del corral). Total: 15 corrales, 10 pollos por corral = 150 pollos.

6. **Aleatorización:** Numerar los 15 corrales del 1 al 15. Generar una permu-tación aleatoria y asignar los primeros 5 al control, los siguientes 5 a dosis baja, los últimos 5 a dosis alta.

7. **Análisis:** ANOVA de un factor (α = 0,05), seguido de prueba de Tukey si F es significativo.

> **Nota:** La UE es el *corral*, *n*o el pollo individual, porque el probiótico se aplica al comedero compartido, de manera que los 10 pollos por corral son unidades de observación y, si se analizara con n = 50 por tratamiento (150 pollos), se cometería pseudorreplicación.

1.5 Principios fundamentales del diseño experimental

Todo diseño experimental se sostiene sobre tres principios formulados por Fisher: **réplica**, **aleatorización** y **control local** (bloqueo). Estos principios no son opcionales ni independientes entre sí; actúan de forma sinérgica para garantizar la validez de las

conclusiones (Fisher, 1935; Montgomery, 2017). Comprender cada uno de ellos, y especialmente su interacción, es la primera habilidad que debe dominar el investigador veterinario.

1.5.1 Réplica

Concepto Clave

La **réplica** es la repetición independiente de cada tratamiento en más de una unidad experimental. Replicar permite: (1) estimar el error experimental, (2) aumentar la precisión de las estimaciones, y (3) ampliar la base de inferencia (Kuehl, 2000).

La réplica no es lo mismo que la *repetición de mediciones* sobre la misma unidad. Si se mide tres veces la temperatura corporal de un mismo bovino, se tienen tres mediciones, pero una sola réplica. Si se mide la temperatura de tres bovinos diferentes sometidos al mismo tratamiento, se tienen tres réplicas. Esta distinción es crucial: las mediciones repetidas en la misma unidad informan sobre la precisión del instrumento de medición, pero las réplicas informan sobre la variabilidad entre unidades experimentales, que es exactamente lo que necesitamos para estimar el error experimental (Oehlert, 2010; Steel et al., 1997).

¿Cuántas réplicas son necesarias? Esta pregunta tiene implicaciones prácticas y éticas. Muy pocas réplicas producen un experimento con baja potencia estadística: la probabilidad de detectar una diferencia real es pequeña, y el experimento habrá desperdiciado animales sin producir resultados concluyentes. Demasiadas réplicas usan más animales de los necesarios, violando el

principio de Reducción de las 3Rs. El número óptimo depende de cuatro factores: (1) la magnitud de la diferencia que se desea detectar (tamaño del efecto), (2) la variabilidad esperada (error experimental), (3) el nivel de significancia (α), y (4) la potencia deseada ($1 - \beta$) (Cohen, 1988; Dell et al., 2002).

Fórmula Clave

Método de la ecuación de recursos para estimar réplicas en estudios con animales (Arifin & Zahiruddin, 2017; Festing, 2018):

El total de unidades experimentales (N) debe satisfacer:

$$E = N - T \geq 10 \text{ y } E \leq 20 \qquad (1.1)$$

Donde E = grados de libertad del error y T = número total de grupos (trata- mientos). Para un DCA con a tratamientos y r réplicas: $N = ar$ y $E = a\,(r - 1)$.

Ejemplo: Con $a = 4$ tratamientos, ¿cuántas réplicas?

$r = 3$: $E = 4(3 - 1) = 8 \rightarrow$ insuficiente ($E < 10$)

$r = 4$: $E = 4(4 - 1) = 12 \rightarrow$ adecuado ($10 \leq E \leq 20$)

$r = 5$: $E = 4(5 - 1) = 16 \rightarrow$ adecuado

$r = 6$: $E = 4(6 - 1) = 20 \rightarrow$ máximo recomendado

$r = 7$: $E = 4(7 - 1) = 24 \rightarrow$ excesivo ($E > 20$, innecesario)

Regla práctica: con 3–4 tratamientos, usar 4–6 réplicas; con 5–6 tratamientos, usar 3–4 réplicas (Festing, 2018).

Aplicación Veterinaria

*E*n un ensayo de evaluación de un antiinflamatorio no esteroidal (AI*N*E) en caninos posquirúrgicos, se asignan 8 perros al AINE y 8 al placebo, de tal manera que cada perro constituye una réplica independiente dentro de su grupo de tratamiento. Si solo se usara 1 perro por tratamiento, resultaría imposible separar el efecto del AINE del efecto atribuible a la variabilidad individual, mientras que con 8 réplicas por tratamiento (N = 16, E = 14) el investigador dispone de suficientes grados de libertad para estimar el error y detectar una diferencia clínicamente relevante, cumpliendo además con el principio de Reducción al no usar más animales de los necesarios (Altman, 1980; Charan & Kantharia, 2013).

1.5.2 Aleatorización

Concepto Clave

La **aleatorización** es la asignación al azar de los tratamientos a las unidades experimentales. Garantiza que cada unidad experimental tenga la misma proba-bilidad de recibir cualquier tratamiento, lo que protege contra sesgos sistemáticos
, tanto conocidos como desconocidos, y valida el uso de pruebas estadísticas inferenciales (Fisher, 1935; Montgomery, 2017).

La aleatorización cumple dos funciones fundamentales que ninguna otra técnica puede reemplazar:

4. **Eliminación de sesgo de selección:** Si el investigador asigna tratamientos según su criterio (“los animales más sanos al tratamiento experimental”), introduce un sesgo que invalida la comparación. La

aleatorización impide esta interferencia, consciente o inconsciente.

5. **Validación de la inferencia estadística:** Las pruebas de hipótesis (prueba *t*, ANOVA) asumen que las observaciones son independientes y que los residuos son aleatorios. La aleatorización es la *única* justificación lógica de estos supuestos. Sin

ella, el valor *p* y los intervalos de confianza carecen de significado formal (Casella, 2008; Fisher, 1935).

Métodos prácticos de aleatorización:

- Tablas de números aleatorios (método clásico): se asigna un número a cada UE y se seleccionan números de la tabla.
- Generadores de números pseudoaleatorios (software: R, Excel, Python): el método más utilizado actualmente. En Excel: =ALEATORIO (). En R: sample (1: n).
- Sorteo simple (sobre, moneda) para diseños pequeños con pocos tratamientos.

Ejemplo Resuelto

Aleatorización paso a paso para un DCA.

Situación: Asignar 3 tratamientos (T1, T2, T3) a 12 corrales (4 réplicas por tratamiento).

Procedimiento:

1. Numerar los corrales del 1 al 12.

2. Generar una permutación aleatoria: supongamos que obtenemos la secuencia 7, 3, 11, 1, 9, 5, 12, 4, 8, 2, 10, 6.

3. Asignar: primeros 4 números → T1, siguientes 4 → T2, últimos 4 → T3.

Tratamiento	Corrales asignados
T1	Corral 7 Corral 3 Corral 11 Corral 1
T2	Corral 9 Corral 5 Corral 12 Corral 4
T3	Corral 8 Corral 2 Corral 10 Corral 6

Verificación: cada corral recibió exactamente un tratamiento; cada tratamiento tiene 4 réplicas; la asignación fue completamente aleatoria.

1.5.3 Control local (bloqueo)

Alerta Importante

La aleatorización no es lo mismo que la haphazardización. Asignar tratamientos "al criterio del investigador" o "alternando uno y uno" no es aleatorización y puede introducir sesgo de selección. Métodos comunes pero incorrectos:

Asignar los primeros animales que llegan al tratamiento A y los siguientes al B (confunde orden de llegada con tratamiento).

Asignar tratamientos alternando izquierda-derecha en un galpón (confunde posición con tratamiento).

Dejar que el técnico "escoja" qué animales van a cada grupo (sesgo consciente o inconsciente).

La aleatorización debe ser verificable: el investigador debe poder documentar el método usado y, idealmente, la semilla del generador aleatorio (Percie Percie du Sert et al., 2020; Lazic, 2016).

Concepto Clave

El **control local** o **bloqueo** consiste en agrupar las unidades experimentales en bloques homogéneos antes de asignar los tratamientos, de modo que la variabilidad entre bloques se separe del error experimental. Bloquear reduce el error y aumenta la potencia de la prueba (Cochran & Cox, 1957).

El principio del bloqueo se apoya en una idea simple pero poderosa: cuando se sabe que ciertas UE resultan más parecidas entre sí que a otras (por ejemplo, animales de la misma camada, vacas en la misma etapa de lactación, parcelas en la misma zona del campo), se

agrupan las UE similares en bloques y se asignan *todos* los tratamientos dentro de cada bloque, de tal manera que las comparaciones entre tratamientos quedan restringidas a bloques homogéneos y la variabilidad entre bloques se "sustrae" del error experimental (Kuehl, 2000; Mead et al., 2012).

Ejemplos de factores de bloqueo en veterinaria:

- Peso corporal inicial (liviano, mediano, pesado).
- Camada o línea genética.
- Etapa de lactación (temprana, media, tardía).
- Ubicación geográfica (finca A, B, C).
- Día de procesamiento o lote de preparación.
- Edad del animal (joven, adulto, viejo).
- Sexo (macho, hembra) cuando es una fuente conocida de variación, pero no es de interés.

La regla práctica de Fisher es: *"Block what you can; randomize what you cannot"* (bloquea lo que puedas; aleatoriza lo que no puedas controlar), enunciado que sintetiza la relación entre los tres principios: el bloqueo controla la variación *conocida*, en tanto que la aleatorización protege contra la variación *desconocida* y la réplica, a su vez, permite *estimar* la variación residual, tal como recoge la Figura 3.

Aplicación Veterinaria

¿Cuándo bloquear en un ensayo veterinario? La decisión de bloquear depende de si existe una fuente de variación conocida que, de no controlarse, inflará el error experimental, de tal manera que la regla práctica se enuncia así: cuando el investigador puede predecir que ciertos animales responderán diferente a los tratamientos por una característica identificable (peso, edad, genética), debe bloquear por esa característica, mientras que si las UE son razonablemente homogéneas (misma raza, edad, peso), un DCA sin bloqueo resulta suficiente y más simple (Bate & Clark, 2014; Ott & Longnecker, 2016).

El bloqueo reduce el error que la réplica estima

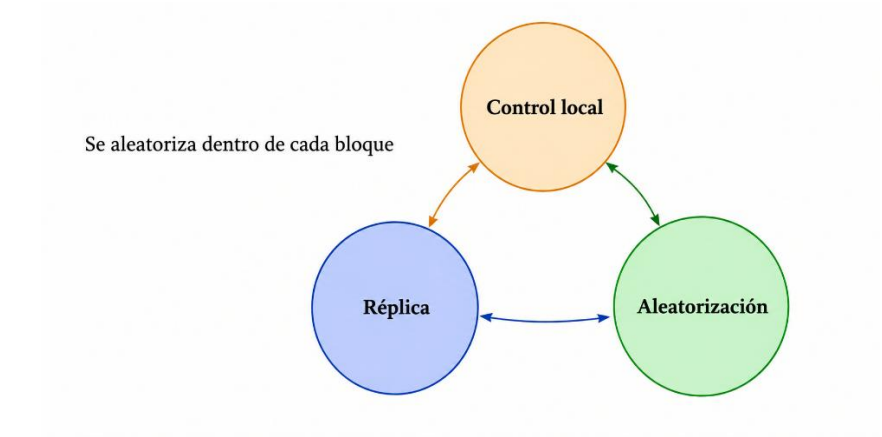


Figura 3. Interrelación entre los tres principios fundamentales del diseño experimental. Cada principio refuerza la función de los otros dos.

La aleatorización valida la estimación del error que la réplica permite.

1.5.4 Interrelación entre los tres principios

Los tres principios operan en conjunto, no de forma independiente. Cada uno refuerza la función de los otros:

Sin réplica no se puede estimar el error experimental; sin aleatorización, la estimación del error está sesgada; sin control local, el error puede ser tan grande que oculte diferencias reales entre tratamientos. Un investigador veterinario que comprende esta tríada está equipado para diseñar experimentos válidos, eficientes y éticos.

1.6 Requisitos básicos de la experimentación

Además de los tres principios, el diseño experimental requiere que el investigador defina con precisión los elementos constitutivos del experimento: la unidad experimental, los tratamientos, la variable respuesta, y las fuentes de error. Una identificación incorrecta de cualquiera de estos elementos compromete la validez de todo el estudio.

1.6.1 La unidad experimental

Concepto Clave

La **unidad experimental** (UE) es la entidad más pequeña a la cual se aplica *independientemente* un tratamiento. Es la unidad de aleatorización y la unidad sobre la cual se calcula el error experimental (Hurlbert, 1984; Kuehl, 2000).

La identificación correcta de la UE constituye posiblemente la decisión

más importante y la más frecuentemente errada en investigación veterinaria, y la regla fundamental es: **la UE es la entidad que recibe el tratamiento de forma independiente**, de tal manera que cuando el tratamiento se aplica a un animal individual (ej.: inyección), la UE es el animal. Si se

aplica a un grupo de animales que comparten un recurso (ej.: dieta en un comedero compartido), la UE es el grupo, no cada animal individual, criterio que la Tabla 1 aplica a los contextos veterinarios más frecuentes.

Tabla 1. Identificación de la unidad experimental en diferentes contextos veterinarios

Contexto	**Tratamiento aplicado a**	**UE**
Fármaco inyectado a cada animal	Animal individual	Animal
Dieta mezclada en comedero de corral	Corral completo	Corral
Vacuna aplicada a cada ternero	Ternero individual	Ternero
Desinfectante aplicado a galpón	Galpón completo	Galpón
Antibiótico en agua de bebedero	Grupo con bebedero compartido	Grupo
Protocolo de ordeño en la sala	Sala de ordeño completa	Sala
Antiparasitario oral a cada oveja	Oveja individual	Oveja
Aditivo en tanque de agua de estanque	Estanque completo	Estanque

1.6.2 Unidad experimental vs. unidad de observación: pseudorreplicación

La **unidad de observación** (UO) es la entidad sobre la cual se registra la medición. La **unidad de muestreo** (UM) es una submuestra dentro de la UO. Es posible, y frecuente, que la UE, la UO y la UM sean entidades diferentes (Figura 4). La confusión entre ellas genera

el error estadístico más grave y prevalente en la literatura veterinaria: la **pseudorreplicación**.

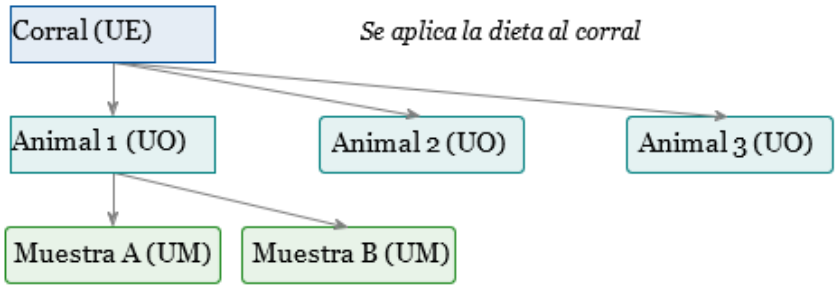


Figura 4. Jerarquía entre unidad experimental (UE), unidad de observación (UO) y unidad de muestreo (UM). Si la dieta se aplica al corral, la UE es el corral, no el animal.

Alerta Importante

Pseudorreplicación ocurre cuando se tratan las unidades de observación como si fueran unidades experimentales, y constituye uno de los errores más frecuentes y graves en investigación veterinaria: Hurlbert (1984) documentó que el 48 % de los estudios ecológicos publicados en revistas de alto impacto cometían este error, Lazic (2010) encontró a su vez tasas similares en neurociencia, y en veterinaria el problema es igualmente prevalente (Festing, 2006).

Para comprender la gravedad de la pseudorreplicación, considérese el siguiente razonamiento: si se asigna la dieta A a un corral con 20 pollos y la dieta B a otro corral con 20 pollos, y se analiza con n = 20 por tratamiento, se están usando 40 observaciones para estimar la variación entre tratamientos, no obstante, la información real sobre esa variación proviene solo de 2 corrales (1 por tratamiento), dado que los 20 pollos dentro de cada corral no son

independientes: comparten el mismo comedero, bebedero, microambiente y manejo. Cualquier diferencia entre corrales (temperatura, corrientes de aire, jerarquía social) se confunde con el efecto de la dieta, y con n = 1 corral por tratamiento **no hay grados de libertad para estimar el error** y no se puede hacer prueba estadística alguna (Lazic, 2016).

Ejemplo Resuelto

Identificación de pseudorreplicación en tres escenarios.

Escenario 1: Un investigador evalúa 2 antiparasitarios en ovinos. Asigna 15 ovejas al fármaco A y 15 ovejas al fármaco B. Cada oveja recibe su dosis individualmente (vía oral). Reporta $n = 15$.

Análisis: La UE es la oveja individual (el tratamiento se aplica individual-mente). $n = 15$ es correcto. **No hay pseudorreplicación.**

Escenario 2: Un nutricionista evalúa 3 dietas en pollos. Usa 3 galpones, asigna una dieta a cada galpón (500 pollos por galpón). Reporta $n = 500$ por tratamiento.

Análisis: La UE es el galpón (la dieta se aplica al comedero comparti-do del galpón). Con 1 galpón por dieta, $n = 1$. Reportar $n = 500$ es pseudorreplicación extrema. **El experimento no tiene réplicas, situación que la Figura 5 contrapone a la asignación correcta.**

Corrección: Usar al menos 4 galpones por dieta (12 galpones total). Calcular el promedio del galpón como la observación. $n = 4$ por

tratamiento.

Escenario 3: U*n* farmacólogo evalúa 2 dosis de un antibiótico, inyecta cada dosis a 10 ratas, toma 3 muestras de sangre de cada rata, analiza cada muestra por duplicado (2 lecturas) y reporta n = 60 por tratamiento (10 ratas × 3 muestras

× 2 lecturas).

Análisis: La UE es la rata, ya que la i*n*yección se aplica individualmente, en tanto que las 3 muestras son submuestreo dentro de ella y las 2 lecturas, repeticiones de la medición, por lo que el dato correcto es el promedio de las 6 lecturas por rata, n = 10, y reportar n = 60 es pseudorreplicación triple. **Corrección: promediar a nivel de UE.**

Caso

Evaluación de un antibiótico en bovinos de engorde.

Un investigador desea evaluar si la oxitetraciclina inyectable reduce la incidencia de enfermedad respiratoria bovina (ERB) en terneros recién llegados a un feedlot, para lo cual dispone de 4 corrales con 25 terneros cada uno y asigna aleatoriamente 2 corrales a oxitetraciclina (inyección metafiláctica al ingreso) y 2 corrales a control (sin tratamiento).
Pregunta: ¿Cuál es la unidad experimental?
Análisis:

CORRECTO

4 corrales por dieta, $n = 4$

INCORRECTO (pseudorreplicación) 1 corral por dieta, reporta *n* = 20

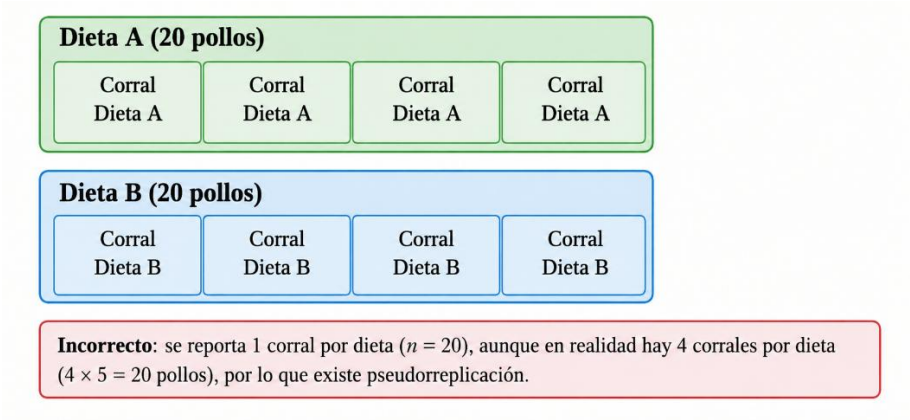


Figura 5. Comparación entre una asignación correcta (izquierda) y pseudorreplicación (derecha) en un ensayo de dietas en aves. A la izquierda, cada corral es una UE independiente; a la derecha, los pollos dentro de un solo corral no son independientes.

- El tratamiento se aplica individualmente a cada ternero (inyección), pero la decisión de tratar es a nivel de corral (metafilaxia de todo el corral).

- Los terneros dentro de un corral comparten ambiente, comedero y bebedero → no son independientes.

- Con n = 2, la potencia es muy baja (E = 2(2 − 1) = 2 del error). Se recomienda aumentar a 4–5 corrales.

1.6.3 Tratamiento y factor

Concepto Clave

Un **tratamiento** es una condición o intervención específica que se aplica a las unidades experimentales, en tanto que un **factor** es una variable independiente cuyos niveles constituyen los tratamientos, y estos factores pueden ser **cualitativos** (tipo de antibiótico: A, B, C) o **cuantitativos** (dosis: 5, 10, 15 mg/kg) (Montgomery, 2017).

En un experimento con un solo factor, cada nivel del factor es un tratamiento, mientras que en un experimento factorial ese papel lo cumple cada *combinación* de niveles de los factores, de tal manera que al evaluar 2 antibióticos × 3 dosis se obtienen 2 × 3 = 6 tratamientos (combinaciones de tratamiento), distinción fundamental para el cálculo correcto de los grados de libertad y la construcción de la tabla ANOVA; los tipos de factor y sus ejemplos veterinarios se recogen en la Tabla 2.

El **tratamiento control** es la condición de referencia contra la cual se comparan los demás tratamientos. Puede ser: (a) un control negativo (sin tratamiento o placebo),

(b) un control positivo (el tratamiento estándar actualmente en uso), o (c) ambos, si bien en investigación veterinaria el uso de placebo debe considerar el bienestar animal, ya que no es ético negar un tratamiento efectivo conocido a animales enfermos sin justificación científica sólida (Festing et al., 2002; National Research Council, 2011).

Tabla 2. Tipos de factores y ejemplos en investigación veterinaria.

Tipo	Característica	Análisis	Ejemplo
Cualitativo	Niveles sin orden numérico	—	Raza: Holstein, Jersey, Brown Swiss
Cuantitativo	Niveles con escala numérica	ANOVA + Tukey	Dosis: 0, 5, 10, 20 mg/kg
Combinación factorial	Combinación de factores	ANOVA + regresión	Mixto: Antibiótico (A, B) × dosis (5, 10)

1.6.4 Variable respuesta

La **variable respuesta** (o variable dependiente) es la característica medida en cada unidad experimental para evaluar el efecto del tratamiento (véanse ejemplos por área en la Tabla 3), y esa elección determina, a su vez, qué información se obtiene del experimento y, en gran medida, su sensibilidad para detectar diferencias, de ahí que una buena variable respuesta deba ser:

a) **Relevante**: directamente relacionada con el objetivo del estudio.

b) **Medible**: cuantificable con un método estandarizado y reproducible.

c) **Sensible**: capaz de detectar diferencias entre tratamientos que sean biológicamente importantes.

d) **Reproducible**: que diferentes evaluadores obtengan resultados similares bajo las mismas condiciones.

Tabla 3. Ejemplos de variables respuesta en investigación veterinaria

Área	Variable	Tipo
Nutrición	Ganancia diaria de peso g/día	Continua
Farmacología	Concentración plasmática µg/mL	Continua
Reproducción	Tasa de concepción %	Proporción
Parasitología	Conteo de OPG huevos/g	Discreta
Patología	Grado de lesión histológica Escala 0–4	Ordinal
Producción	Producción de leche L/día	Continua
Inmunología	Título de anticuerpos $\log_2$	Continua
Microbiología	Recuento de UFC UFC/mL	Discreta

1.6.5 Error experimental

Concepto Clave

El **error experimental** es la variación observada entre unidades experimentales que reciben el mismo tratamiento, variación que refleja la variabilidad biológica natural, la imprecisión de las mediciones y el efecto de factores no controlados, y que constituye, a su vez, la "vara de medir" contra la cual se comparan las diferencias entre tratamientos (Steel et al., 1997).

El error experimental es inevitable: no existe el experimento perfecto con variación cero. Pero sí puede minimizarse mediante tres

estrategias complementarias:

- **Usar material experimental homogéneo:** animales de la misma raza, edad, peso y condición reducen la variabilidad de base.
- **Bloquear por fuentes de variación conocidas:** si la edad de los animales afecta la respuesta, agrupar por edad (bloques) retira esa variación del error.
- **Estandarizar protocolos:** medir siempre con el mismo equipo, a la misma hora, por el mismo evaluador entrenado.
- **Aumentar las réplicas:** más réplicas no reducen el error, pero aumentan la precisión de su estimación y la potencia de la prueba.

En la tabla ANOVA, el error experimental se expresa como el **cuadrado medio del error** (CM_{Error}), de tal manera que un CMError grande indica mucha variación no explicada y un CMError pequeño, buen control experimental, mientras que el estadístico *F* compara la variación entre tratamientos (CM_{Trat}) con la variación del error (CMError): si los tratamientos tienen un efecto real, F será grande (Snedecor & Cochran, 1989).

1.6.6 Hipótesis nula y alternativa

En el contexto del diseño experimental, las hipótesis de investigación se traducen a formulaciones estadísticas que pueden someterse a prueba formal.

Fórmula Clave

Hipótesis nula (H_0): No existe diferencia entre los tratamientos.

(1.2

$H_0: \mu_1 = \mu_2 = \cdots = \mu_a$

Hipótesis alternativa (H_1): Al menos un tratamiento difiere de los demás.

$H_1: \mu_i =$ para al menos un par μ_i (i, i) (1.3

Donde μ_i es la media poblacional del tratamiento i y a es el número total de tratamientos.

La hipótesis nula es la que se somete a prueba estadística, y el experimento busca evidencia para rechazarla, de tal manera que si esa evidencia resulta suficiente (valor $p < \alpha$, donde α es el nivel de significancia predeterminado, típicamente $_0$.05), se rechaza $H0$ y se concluye que existen diferencias significativas, mientras que cuando no lo es ($p\ \alpha$), se retiene H0 $\geq$ —lo que *no* significa que los tratamientos sean iguales, sino que no se encontró evidencia suficiente para afirmar que difieren con el tamaño de muestra utilizado (Pagano & Gauvreau, 2018; Wasserstein & Lazar, 2016).

"**No significativo**" no significa "**sin efecto**". Un resultado p > 0,05 puede deberse a: (a) que realmente no hay diferencia, (b) que la diferencia existe, pero el experimento no tuvo suficiente potencia para detectarla (pocas réplicas, mucha variabilidad), o (c) que la diferencia es pequeña pero real. Confundir "ausencia de evidencia" con "evidencia de ausencia" es un error lógico frecuente. Siempre tamaño del efecto

observado junto con el valor p (Nakagawa & Cuthill, 2007; Watzel, 2016).

1.7 Las variables en la investigación científica

1.7.1 Clasificación de variables

Las variables se clasifican según su función en el experimento, como se observa en la Figura 6, y según su escala de medición. Comprender esta clasificación es esencial para seleccionar las pruebas estadísticas apropiadas y para interpretar correctamente los resultados.

Según su **función** en el experimento:

- **Variable independiente** (factor/tratamiento): la que el investigador manipula deliberadamente. Define los grupos de comparación.

- **Variable dependiente** (respuesta): la que se mide como resultado del efecto del tratamiento. Es lo que el investigador quiere explicar.

- **Variable controlada** (bloqueo): fuente de variación conocida que se controla mediante el diseño (agrupación en bloques) para reducir el error.

- **Variable extraña** (confusora): fuente de variación no controlada que puede sesgar los resultados si no se aleatoriza adecuadamente. Es la variable "oculta" que amenaza la

validez interna del estudio.

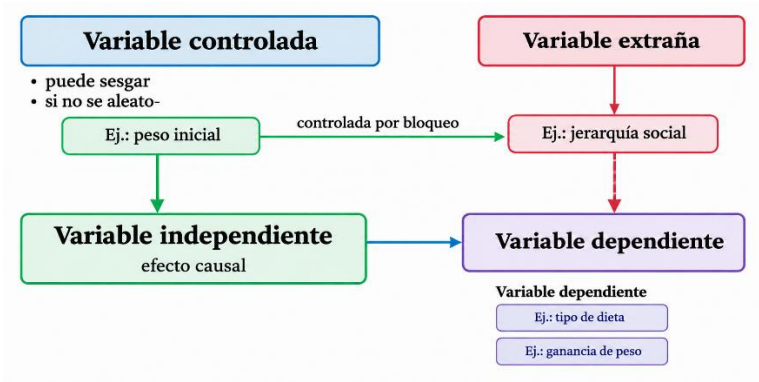


Figura 6. Clasificación de variables según su función en el experimento. Las flechas indican la dirección del efecto y la relación con la variable dependiente.

Ejemplo Resuelto

Identificación de variables en un ensayo veterinario

Situación: Se evalúa en toros Brahman de 3–5 años el efecto de 3 niveles de suplementación con zinc (0, 40 y 80 mg/kg) sobre la calidad del semen, con bloqueo por peso corporal (liviano: < 500 kg, mediano: 500–600 kg, pesado: > 600 kg).

Identificación.

Variable independiente: nivel de zinc (factor cuantitativo con 3 niveles).

Variable dependiente: motilidad espermática (%), concentración (millones/mL), morfología normal (%).

Variable controlada: peso corporal (factor de bloqueo con 3 niveles).

Variables extrañas potenciales: temperatura ambiental, época del año, experiencia de recolección, estrés por manejo → controladas por aleatorización/estandarización de protocolos.

UE: toro individual (la suplementación se administra a cada toro).

Diseño: DBCA con 3 tratamientos y 3 bloques.

1.7.2 Escalas de medición

Tabla 4. Escalas de medición con ejemplos veterinarios y pruebas estadísticas apropia-das.

Escala	Propiedades	Ejemplo	Análisis veterinario
Nominal	Categorías sin orden	Raza: Holstein, Jersey, Angus	Chi-cuadrado
Ordinal	Categorías con orden	Condición corporal: 1, 2, 3, 4, 5	Kruskal-Wallis
Intervalo	Diferencias iguales, sin cero absoluto	Temperatura (°C)	ANOVA, t
Razón	Cero absolutos, proporciones válidas	Peso corporal (kg)	ANOVA, t

La escala de medición de*t*ermina qué pruebas estadísticas son apropiadas (Tabla 4), ya que las variables en escala de razón o intervalo permiten usar pruebas paramétricas (ANOVA, prueba t), en tanto que las ordinales requieren pruebas no paramétricas o tratamientos especiales, y a las nominales les corresponden pruebas

para proporciones o tablas de contingencia (Daniel & Cross, 2013; Zar, 2010). En este libro, el énfasis estará en variables continuas (razón o intervalo), que son las más comunes en investigación experimental veterinaria y las que se analizan con ANOVA.

Concepto Clave

La escala de medición determina las operaciones matemáticas permitidas y, por tanto, las pruebas estadísticas aplicables. Solo las escalas de intervalo y razón permiten calcular medias y varianzas de forma significativa. Un promedio de condiciones corporales (escala ordinal 1–5) es matemáticamente posible pero conceptualmente cuestionable (Petrie & Watson, 2013).

1.8 Medidas de resumen de datos

Antes de aplicar pruebas inferenciales, es fundamental describir los datos. Las medidas de resumen proporcionan una visión concisa de la distribución de los datos: dónde se concentran (tendencia central), cuánto se dispersan (dispersión) y cómo se distribuyen (posición). En el contexto del diseño experimental, estas medidas son los insumos del ANOVA: el ANOVA compara las medias de los tratamientos ($\bar{Y}_i$) y las evalúa contra la variabilidad del error (s^2).

1.8.1 Medidas de tendencia central

Fórmula Clave

Media aritmética ($\bar{Y}$): suma de todos los valores dividida por el número de observaciones.

(1.4) $$\bar{Y} = \frac{1}{n} \sum Y_i$$

Mediana ($\tilde{Y}$): valor central cuando los datos se ordenan de menor a mayor. Si n es par, es el promedio de los dos valores centrales. Es más robusta que la media ante valores atípicos.

Moda: valor que aparece con mayor frecuencia. Puede no ser única o no existir.

La media es la medida más usada en diseño experimental porque el ANOVA compara medias de tratamientos, si bien es sensible a valores atípicos, ya que un solo valor extremo puede desplazarla considerablemente, mientras que la mediana es más robusta ante valores extremos y resulta preferible cuando la distribución es asimétrica o hay outliers, criterio que la Tabla 5 resume para cada medida (Snedecor & Cochran, 1989; Zar, 2010).

Tabla 5. Comparación de las medidas de tendencia central: cuándo usar cada una.

Medida veterinaria	Usar cuando…	Ejemplo
Media	Distribución simétrica, sin outliers	GDP en pollos de engorde

Mediana	Distribución asimétrica u outliers		Conteo de OPG (muy variable)
Moda	Variables discretas o categóricas		Puntaje de claudicación (0–4)

1.8.2 Medidas de dispersión

Fórmula Clave

Varianza muestral (s^2): promedio de las desviaciones cuadráticas respecto a la media.

$$s^2 = \frac{1}{n-1}\sum_{j=1}^{n}(Y_j - \bar{Y})^2 \qquad (1.5)$$

Desviación estándar (s): raíz cuadrada de la varianza, expresada en las mismas unidades que los datos

$$s = \sqrt{s^2} \qquad (1.6)$$

Coeficiente de variación (CV): expresa la dispersión como porcentaje de la media. Permite comparar variabilidad entre variables con diferentes unidades o

Magnitud es.

$$CV = \frac{s}{\bar{Y}} \times 100\% \qquad (1.7)$$

Concepto lave

¿Por qué se divide por $n-1$ y no por n? La varianza muestral usa $n-1$ (grados de libertad) en el denominador porque la media muestral $\bar{Y}$ ya "consume" un grado de libertad. Si dividimos por n, subestimaríamos sistemáticamente la varianza poblacional. Esta corrección, conocida como **corrección de Bessel**, garantiza que s^2 sea un estimador **insesgado** de la varianza poblacional σ^2 (Daniel & Cross, 2013; Snedecor & Cochran, 1989).

Aplicación Veterinaria

En investigación con animales, el CV es especialmente útil para comparar la variabilidad entre variables con diferentes unidades o entre especies con diferentes magnitudes. Orientaciones generales para ensayos con animales de producción:

- CV < 10 %: variabilidad baja, excelente control experimental.
- 10 % ≤ CV ≤ 20 %: variabilidad moderada, aceptable para la mayoría de variables.
- 20 % < CV ≤ 30 %: variabilidad alta, considerar más réplicas o mejor control.
- CV > 30 %: variabilidad muy alta, revisar la homogeneidad del material experimental o considerar transformación de datos.

Estas referencias son orientativas; variables como conteos parasitarios (OPG) comúnmente presentan CV > 50 % y pueden requerir transformación logarítmica o raíz cuadrada antes del análisis (Petrie & Watson, 2013; Thrusfield, 2018).

Ejemplo Resuelto

Cálculo completo de medidas descriptivas con 12 datos.

Se registró el peso al destete (kg) de 12 terneros Brahman:

180	195	172	188	165	198	182	175	190	185	170	200

Paso 1. Media:

$$\bar{y} = \frac{180 + 195 + 172 + 188 + 165 + 198 + 182 + 175 + 190 + 185 + 170 + 200}{12}$$

$$\bar{y} = \frac{2200}{12} = 183{,}33 \text{ kg}$$

Paso 2. Desviaciones cuadráticas:

j	Y_j	$Y_j - \bar{y}$	$(Y_j - \bar{y})^2$	j	Y_j	$Y_j - \bar{y}$	$(Y_j - \bar{y})^2$
1	180	−3,33	11,09	7	182	−1,33	1,77
2	195	11,67	136,19	8	175	−8,33	69,39
3	172	−11,33	128,37	9	190	6,67	44,49
4	188	4,67	21,81	10	185	1,67	2,79
5	165	−18,33	336,00	11	170	−13,33	177,69
6	198	14,67	215,21	12	200	16,67	277,89

Suma de cuadrados = 1422,67

Paso 3. Varianza:

$$s^2 = \frac{1422{,}67}{12 - 1} = \frac{1422{,}67}{11} = 129{,}33 \text{ kg}^2$$

Paso 4. Desviación estándar:

$$s = \sqrt{129{,}33} = 11{,}37 \text{ kg}$$

Paso 5. Coeficiente de variación:

$$CV = \frac{11{,}37}{183{,}33} \times 100\,\% = 6{,}20\,\%$$

Figura 7. Ejemplo Resuelto.

1.8.3 Medidas de posición

Los **cuartiles** dividen los datos ordenados en cuatro partes iguales:

- Q_1 (primer cuartil): el 25 % de los datos está por debajo.
- Q_2 (segundo cuartil): coincide con la mediana.
- Q_3 (tercer cuartil): el 75 % de los datos está por debajo.

Rango intercuartílico (RIC): RIC = Q_3 Q_1, mide la dispersión central del 50 % medio de los datos. Es robusto ante valores extremos.

Los **percentiles** generalizan los cuartiles, dado que el percentil *k* es el valor por debajo del cual cae el k % de los datos, de donde Q_1 = P_{25}, Q_2 = P_{50} y Q_3 = P_{75}, y su empleo resulta útil para establecer rangos de referencia en medicina veterinaria: "el 9_5 % de los animales sanos tienen valores entre P2,5 y P_{97},5" (Daniel & Cross, 2013).

1.8.4 Representación gráfica

Los diagramas de caja (*boxplots*) son la herramienta gráfica estándar en investigación experimental para comparar la distribución de una variable entre tratamientos. Cada caja muestra la mediana, Q_1, Q_3 y los valores atípicos, tal como se aprecia en la Figura 8.

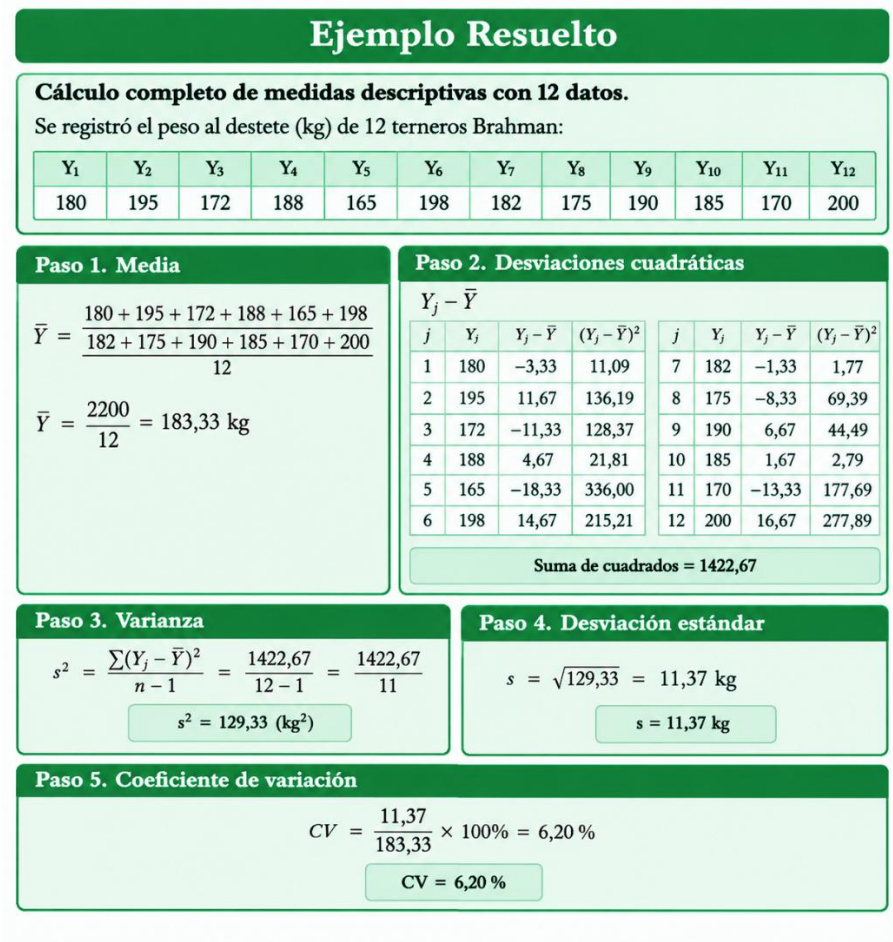


Figura 8. Ejemplo Resuelto 2.

Interpretación: La producción promedio fue 29,34, 2,39 L/día ± con un CV de 8.15 %, indicando variabilidad baja y una muestra bastante homogénea para vacas lecheras de alta producción.

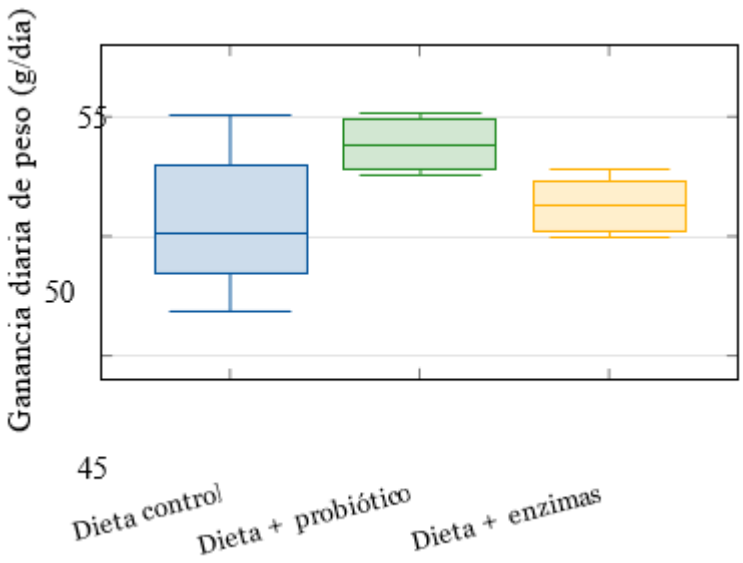


Figura 9.Diagrama de cajas (boxplot) comparativo de tres dietas en pollos de engorde. La caja representa el rango intercuartílico (Q1 a Q3), la línea horizontal dentro de la caja es la mediana, y los bigotes se extienden hasta los valores mínimo y máximo.

1.8.5 Presentación estándar de resultados

En publicaciones científicas, los datos descriptivos de cada grupo se reportan como:

- $\bar{Y} \pm s$ (o $\bar{Y} \pm$ EE, donde EE = $s/\sqrt{n}$ es el error estándar de la media), junto con el número de observaciones n.

Fórmula Clave

Error estándar de la

$$EE = \frac{s}{\sqrt{}} \quad (1.8)$$

El EE mide la precisión de la estimación de la media, no la dispersión de los datos. El EE disminuye al aumentar *n* (más datos = mejor estimación de la media). La DE mide la dispersión de los datos individuales y no depende de *n* (Pagano & Gauvreau, 2018).

Tabla 6. Formato estándar de presentación de resultados descriptivos (ejemplo: ganancia diaria de peso en pollos)

Tratamiento	n	$\bar{Y} \pm s$ (g/día)	CV (%)
Dieta control	8	51,89 ± 2,25	4.34
Dieta + probiótico	8	56,42 ± 2,81	4.98
Dieta + enzimas	8	54,15 ± 3,10	5.72

Alerta Importante

¿DE o EE? Un error frecuente en artículos veterinarios es usar el error estándar (EE) cuando se debería usar la desviación estándar (DE), o viceversa. Regla simple:

- Use **DE** ($\bar{Y} \pm s$) cuando quiera describir la *dispersión* de los datos individuale
- Use **EE** ($\bar{Y} \pm$ EE) cuando quiera mostrar la *precisión* de la estimación de la media

El EE siempre es menor que la DE (por factor $\sqrt{n}$), de modo que emplear el EE "hace que los datos se vean menos variables", razón por la cual algunos autores lo utilizan incorrectamente para que sus gráficos se vean "más limpios" (Cumming, **2009).**

En tablas descriptivas, prefiera la DE, con el formato que recoge la Tabla 6. En gráficos de barras con comparaciones, puede usar EE, pero declárelo claramente en la leyenda.

1.9 Conceptos clave

- Un **experimento** requiere manipulación deliberada, control y aleatorización. Sin estos tres elementos, es un estudio observacional.
- Los tres principios fundamentales del diseño experimental son **réplica**, **aleatori-zación** y **control local** (bloqueo). Operan sinérgicamente.
- La **unidad experimental** es la entidad a la cual se aplica

independientemente el tratamiento; confundirla con la unidad de observación genera pseudorreplicación.

- La **pseudorreplicación** invalida la prueba estadística al inflar artificialmente los grados de libertad del error.

- Las **variables** se clasifican por función (independiente, dependiente, confusora) y por escala de medición (nominal, ordinal, intervalo, razón).

- El **coeficiente de variación** permite comparar la dispersión entre variables con diferentes unidades.

- Las hipótesis estadísticas (H_0 y H_1) formalizan la pregunta de investigación en términos que pueden probarse con datos.

- El **error estándar** mide la precisión de la media; la **desviación estándar** mide la dispersión de los datos.

- El **método de la ecuación de recursos** permite estimar el número mínimo de réplicas necesarias.

- Planificar el análisis estadístico *antes* de recolectar datos previene el *p-hacking*.

1.10 Actividades de consolidación

- **Actividad 1:** Seleccione un artículo de investigación veterinaria publicado en una revista indexada e identifique en él: (a) el diseño experimental, (b) la unidad experimental, (c) los factores y sus niveles, (d) la variable respuesta, (e) el

número de réplicas, y determine si existe riesgo de pseudorreplicación.

- **Actividad 2:** Proponga un experimento para evaluar el efecto de tres niveles de suplementación con vitamina E (0, 200, 400 UI/día) sobre la calidad seminal en toros reproductores, definiendo la UE, el factor, los niveles y la variable respuesta, junto con un esquema de aleatorización, además de justificar el número de réplicas mediante el método de la ecuación de recursos.

- **Actividad 3:** Con los siguientes datos de producción de leche (L/día) de 10 vacas Holstein, calcule media, varianza, desviación estándar, CV, error estándar y los tres cuartiles: 28.5, 31.2, 26.8, 30.1, 27.4, 33.0, 29.7, 25.9, 32.5, 28.3. Interprete el CV en contexto de producción lechera.

- **Actividad 4:** Explique por qué es incorrecta la siguiente afirmación: "Administramos el tratamiento a 20 ratas en la misma jaula y obtuvimos 20 réplicas", y proponga un rediseño que resuelva el problema, calculando cuántas jaulas necesitaría para tener al menos 10 GL del error.

1.11 Ejercicios propuestos

1. Un investigador evalúa 4 protocolos de analgesia posquirúrgica en gatos. Asigna 6 gatos a cada protocolo. Identifique: factor, niveles, UE, variable respuesta potencial. Calcule los GL del error si se usa un DCA.

2. E*n* un ensayo de fertilización de pasturas se aplican 3 tipos de fertilizante a parcelas de 100 m^2, cada una de las cuales contiene 5 puntos de muestreo donde se corta y pesa el forraje. ¿Cuál es la UE, cuántas réplicas tiene el ensayo si hay 4 parcelas por fertilizante y cuál es el n correcto para el ANOVA?

3. Clasifique las siguientes variables según su escala de medición:

- Raza del animal,
- Condición corporal (1–5), (c) peso al destete, (d) presencia/ausencia de mastitis,
- (e) temperatura rectal, (f) conteo de células somáticas. Para cada una, indique qué tipo de prueba estadística sería apropiada.

4. Se registraron los pesos al nacimiento (kg) de 12 terneros: 32, 28, 35, 30, 33, 27, 31, 34, 29, 36, 30, 32. Calcule $\bar{Y}$, s^2, s, CV, mediana, Q_1, Q_3 y RIC. *Respuesta:* $\bar{Y}$ = 31,42, s = 2,78, $CV = 8.85\%$.

5. Se mide la concentración plasmática de un fármaco (μg/mL) en 6 perros: 4.2, 3.8, 5.1, 4.5, 3.9, 4.7. Calcule la media, la desviación estándar y el error estándar.

Respuesta: $\bar{Y}$ = 4,37, s = 0,48, EE = 0,20.

6. Una investigación evalúa el efecto de un antiparasitario en ovinos, en la cual se mide la reducción del conteo de huevos por gramo de heces (OPG) antes y después del tratamiento en

15 ovejas. ¿Se trata de un diseño experimental o de un estudio observacional, cuál es la variable independiente y qué tipo de prueba estadística sería apropiada?

7. ¿Por qué se divide la suma de cuadrados por $n - 1$ (y no por n) para calcular la varianza muestral? Explique conceptualmente el concepto de grados de libertad.

8. Un ensayo clínico en equinos evalúa dos antiinflamatorios (meloxicam vs. fenilbu-tazona) y un placebo. Se asignan 10 caballos a cada grupo. Formule H_0 y H_1 en términos estadísticos. ¿Cuántos GL del error tiene el ANOVA?

9. ¿Cuál es la diferencia entre replicar (tener 5 animales por tratamiento) y repetir mediciones (medir 5 veces el mismo animal)? ¿Cuál de las dos contribuye a estimar el error experimental? ¿Pueden combinarse ambas estrategias?

10. Un investigador afirma que "aleatoricé" los tratamientos asignando los animales más pesados al grupo A y los más livianos al grupo B. ¿Por qué esta afirmación es incorrecta? ¿Qué sesgo podría introducir?

11. Deter*m*ine, mediante el método de la ecuación de recursos, cuántas réplicas necesita un experimento con 5 tratamientos en un DCA para tener entre 10 y 20 GL del error. *Respuesta:* r = 3 *da E* = 10*;* r = 5 da E = 20. *Rango recomendado: 3–5 réplicas.*

12. Un estudio evalúa 3 vacunas contra leptospirosis canina. Se

usa un DCA con 8 perros por vaeuna. Calcule los GL del error. Si el investigador decide bloquear por edad (cachorro, adulto, geronte), ¿cómo cambian los GL del error? ¿Cuántos bloques tiene? *Respuesta: DCA:* $E = 3(81) = 21$. *DBCA: 3 bloques* $3 \text{ tratamientos} = 9 \text{ perros; } E = (31)(31) = 4$ *— necesitaría más perros por bloque.*

1.12 Preguntas de autoevaluación

Selección múltiple (seleccione la opción correcta):

1. ¿Cuál es la principal diferencia entre un experimento y un estudio observacional?

a) El tamaño de muestra.

b) La manipulación deliberada del factor de interés. ←

c) El uso de estadística inferencial.

d) La publicación en revista indexada.

2. La pseudorreplicación ocurre cuando:

a) Se usan demasiadas réplicas.

b) Se confunde la UO con la UE. ←

c) Se aleatoriza incorrectamente.

d) Se omite el grupo control.

3. El principal objetivo de la aleatorización es:

a) Aumentar el tamaño de muestra.

b) Evitar sesgos sistemáticos en la asignación de tratamientos. ←

c) Reducir el costo del experimento.

d) Simplificar el análisis estadístico.

4. El bloqueo es recomendable cuando:

a) Las unidades experimentales son homogéneas.

b) Existe una fuente de variación conocida que se quiere controlar. ←

c) Se dispone de pocos tratamientos.

d) Se trabaja con variables nominales.

5. Un CV de 8 % en ganancia de peso de pollos indica:

a) Variabilidad inaceptable.

b) Variabilidad baja y buen control experimental. ←

c) Que se necesitan más réplicas.

d) Que los datos no son normales.

6. Si el tratamiento se aplica al agua del bebedero compartido por un grupo de aves, la UE es:

a) Cada ave individual.

b) El bebedero.

c) El grupo de aves que comparte el bebedero. ←

d) El galpón.

7. La hipótesis nula (H_0) en un ANOVA afirma que:

a) Todos los tratamientos son diferentes.

b) Al menos un tratamiento es diferente.

c) Todas las medias poblacionales son iguales. ←

d) La varianza es cero.

8. La desviación estándar se expresa en:

a) Las mismas unidades que la varianza.

b) Las mismas unidades que los datos originales. ←

c) Porcentaje.

d) Unidades estandarizadas.

9. El método de la ecuación de recursos indica que los GL del error deben estar entre:

a) 5 y 10.

b) 10 y 20. ←

c) 20 y 50.

d) No hay límites definidos.

10. ¿Cuál de las siguientes NO es una función de la réplica?

a) Estimar el error experimental.

b) Aumentar la precisión de las estimaciones.

c) Eliminar sesgos de selección. ←

d) Ampliar la base de inferencia.

11. El error estándar de la media (EE) se diferencia de la desviación estándar (DE) en que:

a) El EE mide la precisión de la estimación de la media, la DE mide la dispersión de los datos. ←

b) Son lo mismo, pero con diferente nombre.

c) El EE siempre es mayor que la DE.

d) El EE no depende del tamaño de muestra.

12. En un DCA con 5 tratamientos y 4 réplicas, los GL del error son:

a) 15. ←

b) 19.

c) 20.

d) 4.

Preguntas de desarrollo:

1. Explique con un ejemplo veterinario la diferencia entre un factor cualitativo y un factor cuantitativo. ¿Cómo afecta esta distinción al análisis estadístico posterior?

2. Describa un escenario de investigación en producción lechera donde el bloqueo por etapa de lactación sea apropiado. ¿Qué pasaría si no se bloquea? ¿Cómo se estima si el bloqueo fue eficiente?

3. Fre*n*te al experimento que le presenta un colega, en el que evaluó 3 dietas en cerdos empleando 3 corrales, cada uno con 10 cerdos, y reporta n = 30 réplicas, escriba una crítica constructiva que explique el error y cómo debería rediseñarse el ensayo, incluyendo el cálculo del número real de réplicas y los GL del error.

1.13 Glosario

Aleatorización Asignación al azar de tratamientos a unidades experimentales para evitar sesgos sistemáticos y validar la inferencia estadística.

Bloqueo (control local)

Agrupación de UE homogéneas en bloques para reducir el error experimental separando la variabilidad entre bloques.

Coeficiente de variación (CV)

Medida de dispersión relativa expresada como porcentaje de la media (CV = $s/\bar{Y}$ × 100 %). Permite comparar variabilidad entre variables con diferentes unidades.

Confusión

Situación en la que el efecto de un factor de interés no se puede separar del efecto de otra variable no controlada.

Desviación estándar (s)

Raíz cuadrada de la varianza; expresa la dispersión en las unidades originales de los datos.

Ecuación de recursos

Método para estimar el número de réplicas basado en los grados de libertad del error: $E = N - T$, donde $10 \leq E \leq 20$.

Error estándar (EE)

Desviación estándar de la media muestral ($EE = s/\sqrt{n}$); mide la precisión de la estimación de la media.

Error experimental

Variación entre UE tratadas idénticamente; base para la inferencia estadística.

Escala de intervalo

Escala con diferencias iguales, pero sin cero absoluto (ej.: tempera-tura en ˚C).

Escala de razón

Escala con cero absoluto y proporciones significativas (ej.: peso en kg).

Escala nominal

Escala de categorías sin orden (ej.: raza, color).

Escala ordinal

Escala de categorías con orden, pero intervalos desiguales (ej.: condición corporal 1–5).

Experimento

Investigación con manipulación deliberada de factores, control de condiciones y aleatorización.

Factor

Variable independiente cuyos niveles definen los tratamientos.

Hipótesis alternativa (H_1)

Afirmación de que existe al menos una diferencia entre tratamientos.

Hipótesis nula (H0)

Afirmación de que no existen diferencias entre tratamientos; es la

que se somete a prueba estadística.

Media aritmética ($\bar{Y}$)

Suma de valores dividida por el número de observaciones.

Potencia estadística ($1-\beta$)

Probabilidad de rechazar H_0 cuando esta es falsa (detectar un efecto real).

Pseudorreplicación

Error que consiste en tratar UO como UE, inflando artificialmente los grados de libertad y la significancia.

Réplica

Repetición independiente de un tratamiento en diferentes UE; no confundir con repetición de mediciones.

Tres Rs (3Rs)

Principios éticos para investigación con animales: Reemplazo, Reducción y Refinamiento.

Tratamiento

Condición experimental específica aplicada a las UE.

Tratamiento control

Tratamiento de referencia (placebo o estándar) contra el cual se

comparan los tratamientos experimentales.

Unidad de muestreo (UM)

Submuestra dentro de la unidad de observación.

Unidad de observación (UO)

Entidad sobre la que se registra la medición; puede ser una subdivisión de la UE.

Unidad experimental (UE)

Entidad más pequeña a la cual se aplica independientemente un tratamiento.

Variable confusora

Variable extraña que afecta la respuesta y está correlacionada con el tratamiento, amenazando la validez interna.

Variable dependiente

Variable medida como respuesta al efecto del tratamiento.

Variable independiente

Variable manipulada por el investigador (factor/tratamiento).

Varianza (s^2)

Promedio de las desviaciones cuadráticas respecto a la media, dividido por $n - 1$.

1.14 Resumen

El diseño experimental es la disciplina que permite planificar investigaciones de modo que las conclusiones resulten válidas, eficientes y éticamente responsables, y se sustenta en tres principios fundamentales establecidos por Fisher (1935): la **réplica**, que permite estimar el error experimental y constituye la base de la inferencia estadística; la **aleatorización**, que protege contra sesgos sistemáticos, tanto conocidos como desconocidos, y valida

el uso de pruebas estadísticas; y el **control local** (bloqueo), que reduce la variabilidad no deseada agrupando unidades experimentales homogéneas.

La identificación correcta de la **unidad experimental** es tal vez el concepto más crí-tico y el más frecuentemente violado en investigación veterinaria. La pseudorreplicación

(tratar unidades de observación como si fueran unidades experimentales) conduce a conclusiones falsas al inflar los grados de libertad del error, equivocación documentada por Hurlbert (1984) en ecología y por Lazic (2010) en neurociencia, que resulta igualmente prevalente en la literatura veterinaria y constituye una amenaza seria a la validez de los hallazgos publicados.

Las **variables** se clasifican según su función (independiente, dependiente, confusora) y su escala de medición (nominal, ordinal, intervalo, razón), lo que determina las pruebas estadísticas apropiadas, y a su vez las **medidas descriptivas**, media, varianza, desviación

estándar y coeficiente de variación, proporcionan un resumen cuantitativo de los datos antes del análisis inferencial y constituyen los insumos fundamentales del ANOVA. El coeficiente de variación resulta particularmente útil en investigación veterinaria para evaluar la homogeneidad del material experimental y comparar variabilidades entre diferentes variables o especies.

La formulación precisa de hipótesis estadísticas (H_0 y H_1) traduce la pregunta de investigación a términos que pueden probarse con datos, en tanto que la determinación del número adecuado de réplicas, ni demasiadas ni demasiado pocas, requiere considerar el tamaño del efecto esperado, la variabilidad del material experimental, el nivel de significancia y la potencia deseada, respetando siempre los principios éticos de uso de animales en investigación (Festing, 2018; Russell & Burch, 1959).

En el próximo capítulo se formalizará el proceso de prueba de hipótesis mediante la prueba *t* de Student y el análisis de varianza (ANOVA), que constituyen las herramientas inferenciales centrales del diseño experimental.

CAPÍTULO II

2 PRUEBA DE HIPÓTESIS Y ANÁLISIS DE VARIANZA

En el capítulo anterior se establecieron los fundamentos conceptuales del diseño experimental: los principios de réplica, aleatorización y control local, la identificación de la unidad experimental y las herramientas descriptivas básicas. Con esas bases, estamos preparados para abordar la maquinaria inferencial que permite convertir datos en decisiones: la prueba de hipótesis y el análisis de varianza (ANOVA).

2.1 Introducción

La estadística descriptiva resume los datos, pero no responde la pregunta central de un experimento: ¿las diferencias observadas entre tratamientos son reales o podrían deberse al azar? Para responder esta pregunta se necesita la **inferencia estadística**, cuya herramienta principal en el diseño experimental es el **análisis de varianza** (ANOVA). El ANOVA, desarrollado por Fisher (1925) en la Estación Experimental de Rothamsted, revolucionó la investigación empírica al proporcionar un marco lógico para evaluar simultáneamente las diferencias entre múltiples tratamientos (Montgomery, 2017).

La transición de la estadística descriptiva a la inferencial representa un salto conceptual importante, ya que las herramientas del Capítulo 1, media, varianza y coeficiente de variación, describen el comportamiento de los datos, aunque no permiten afirmar con

confianza que las diferencias observadas sean reales y no producto del azar muestral. Un grupo de pollos con probiótico puede mostrar una media de ganancia de peso mayor que el grupo control, pero ¿es esa diferencia lo suficientemente grande para concluir que el probiótico funciona, o podría haberse obtenido una diferencia similar simplemente por variación aleatoria entre animales? Ante esta situación, la prueba de hipótesis formaliza este razonamiento (Daniel & Cross, 2013; Pagano & Gauvreau, 2018).

Este capítulo desarrolla el marco completo de la prueba de hipótesis, desde los conceptos de error tipo I y tipo II hasta el procedimiento formal de cinco pasos, y luego presenta el ANOVA como extensión natural de la prueba t para comparar más de dos tratamientos, a su vez, se incluyen las pruebas de comparaciones múltiples (Tukey) para identificar *cuáles* tratamientos difieren, y se ilustra todo el proceso con ejemplos integradores de investigación veterinaria resueltos paso a paso con toda la aritmética explícita.

2.2 Objetivos de aprendizaje

1. Definir hipótesis nula y alternativa y diferenciar los errores tipo I y tipo II.

2. Interpretar correctamente el valor p y el nivel de significancia α.

3. Aplicar el procedimiento de cinco pasos para la prueba de hipótesis.

4. Ejecutar una prueba t de Student para dos muestras

independientes y para muestras pareadas con datos veterinarios.

5. Construir una tabla ANOVA completa y calcular el estadístico F paso a paso, mostrando toda la aritmética.

6. Aplicar la prueba de Tukey (HSD) para identificar diferencias entre pares de tratamientos y asignar letras de agrupamiento.

7. Verificar los supuestos del ANOVA e identificar cuándo se violan.

8. Reportar resultados de ANOVA en formato APA.

2.3 Preguntas de enfoque

1. ¿Por qué no se puede usar la prueba *t* repetidamente para comparar más de dos tratamientos?

2. ¿Qué significa realmente "rechazar H_0 con $p < 0,05$"?

3. ¿Cómo descompone el ANOVA la variabilidad total de los datos para detectar diferencias entre tratamientos?

4. Si el ANOVA indica diferencias significativas, ¿cómo se determina cuáles trata-mientos difieren entre sí?

5. ¿Cuáles son los supuestos del ANOVA y qué ocurre si se violan?

2.4 Prueba de hipótesis

La prueba de hipótesis es el procedimiento formal que permite al

investigador tomar decisiones basadas en datos. No es un acto mecánico de comparar números con tablas, sino un marco lógico de razonamiento bajo incertidumbre. Antes de calcular cualquier estadístico, el investigador debe comprender qué está preguntando, qué riesgos acepta, y qué significan los posibles resultados.

2.4.1 Concepto de hipótesis estadística

Una **hipótesis estadística** es una afirmación formal sobre un parámetro poblacional que puede someterse a prueba con datos muestrales. En el contexto experimental, siempre se trabaja con un par de hipótesis complementarias que abarcan todas las posibilidades: la hipótesis nula y la alternativa.

Fórmula Clave

Hipótesis nula (H_0): afirma que no hay efecto o diferencia. Es la hipótesis conservadora, la que asume el "statu quo."

Hipótesis alternativa (H_1 o H_a): afirma que sí hay efecto o diferencia. Es lo que el investigador espera demostrar.

Para comparar dos medias: (bilateral)

$$H_0 : \mu_1 = \mu_2 \quad (2.1)$$

$$H_1 : \mu_1 = \mu_2 \quad (2.2)$$

Para comparar a medias

$$H_0 : \mu_1 = \mu_2 = \cdot \quad (2.3)$$

$$\cdot\cdot = \mu_a \quad (2.4)$$

H_1 : al menos un par $\mu_i \neq \mu_i$

La prueba de hipótesis opera bajo una lógica de "reducción al absurdo": se asume temporalmente que H_0 es verdadera y se pregunta cuán probable es obtener los datos observados bajo esa suposición, de tal manera que, si los datos son muy improbables bajo H0 (probabilidad menor que α), se concluye que H0 probablemente es falsa y se rechaza. Este proceso *no demuestra* que H0 sea verdadera ni que H_1 sea verdadera; solo evalúa la evidencia en contra de H0 (Pagano & Gauvreau, 2018; Wasserstein & Lazar, 2016).

Es importante distinguir entre hipótesis **unilaterales** y **bilaterales**. La hipótesis bilateral (H_1: $\mu_1 = \mu_2$) solo pregunta si los tratamientos difieren, sin especificar la dirección. La hipótesis unilateral (H_1: $\mu_1 > \mu_2$) predice la dirección del efecto. En diseño experimental veterinario, generalmente se usan pruebas bilaterales salvo que exista una justificación teórica fuerte para predecir la dirección del efecto *antes* de recolectar datos (Zar, 2010).

2.4.2 Errores tipo I y tipo II

Al tomar una decisión estadística basada en datos muestrales, se pueden cometer dos tipos de errores, cuyas consecuencias confronta la Tabla 7. Estos errores son inherentes al proceso inferencial y no se pueden eliminar simultáneamente; solo se pueden controlar.

Concepto Clave

Error tipo I (α): rechazar H_0 cuando es verdadera ("falso positivo"), lo que implica declarar que un tratamiento funciona cuando en realidad no tiene efecto, y su probabilidad de ocurrencia se controla con el **nivel de significancia** α, convencionalmente fijado en 0.05 (5 %).

Error tipo II (β): no rechazar H_0 cuando es falsa ("falso negativo"). Conse-cuencia: declarar que un tratamiento no funciona cuando en realidad sí tiene efecto. La probabilidad complementaria $1 - \beta$ se denomina **potencia estadística** (Cohen, 1988).

Tabla 7. Tipos de error en la prueba de hipótesis y sus consecuencias.

Decisión	H_0 verdadera	H_0 falsa
No rechazar H_0	Decisión correcta ($1 - \alpha$)	Error tipo II (β)
Rechazar H_0	Error tipo I (α)	Decisión correcta ($1 - \beta$)

Existe una relación inversa entre α y β para un tamaño de muestra fijo: al reducir α (hacerlo más estricto, por ejemplo 0.01), se reduce la probabilidad de falsos positivos, pero se aumenta la probabilidad de falsos negativos (β sube), de tal manera que la única forma de reducir ambos errores simultáneamente es aumentar el tamaño de muestra (Ott & Longnecker, 2016).

Aplicación Veterinaria

Consecuencias prácticas de los errores en veterinaria:

Error tipo I (falso positivo): un AINE es declarado "efectivo" para el dolor posquirúrgico en caninos cuando en realidad no lo es, de tal manera que los veterinarios lo prescriben innecesariamente, los animales quedan sin analgesia adecuada y el fármaco ineficaz desplaza a opciones mejores.

Error tipo II (falso negativo): Un antibiótico realmente efectivo contra mastitis bovina es declarado "sin efecto" porque el ensayo tenía pocas réplicas. Consecuencia: un tratamiento útil no se adopta, y los animales siguen recibiendo protocolos subóptimos.

En investigación veterinaria, se acepta $\alpha = 0{,}05$ y se busca una potencia de al menos $1 - \beta = 0{,}80$ (80 %) (Altman, 1980; Dell et al., 2002).

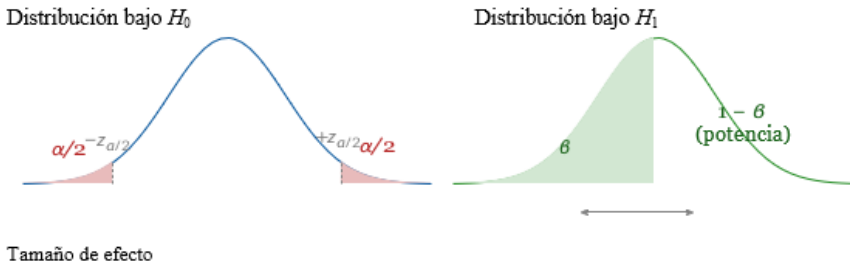


Figura 10.Relación entre errores tipo I (α), tipo II (β) y potencia (1 β). La zona sombreada roja es la región de rechazo bajo H0. La zona verde clara es β (probabilidad de no detectar un efecto real). La zona verde oscura es la potencia.

2.4.3 Nivel de significancia y valor p

Concepto Clave

El **valor** p es la probabilidad de obtener un resultado tan extremo o más extremo que el observado, *asumiendo que* H_0 *es verdadera*, lo cual no equivale a la probabilidad de que H0 sea verdadera ni a la probabilidad de que el resultado se deba al azar, sino a una medida de la **incompatibilidad** de los datos con H0 (Wasserstein & Lazar, 2016).

Regla de decisión: Si $p \leq \alpha$, se rechaza H_0. Si $p > \alpha$, no se rechaza H_0.

Alerta Importante

Errores comunes en la interpretación del valor p:

- **Incorrecto:** "$p = 0{,}03$ significa que hay un 3 % de probabilidad de que el resultado se deba al azar."
- **Correcto:** "Si H_0 fuera verdadera, la probabilidad de observar una diferencia tan grande o mayor es 0.03."
- **Incorrecto:** "$p > 0{,}05$ demuestra que no hay diferencia."
- **Correcto:** "No se encontró evidencia suficiente para rechazar H_0 al nivel $\alpha = 0{,}05$."
- **Incorrecto:** "$p = 0{,}001$ indica un efecto grande."
- **Correcto:** "$p = 0{,}001$ indica fuerte evidencia contra H_0, pero el tamaño del efecto puede ser pequeño si n es grande."

La significancia estadística ($p < 0{,}05$) no implica necesariamente relevancia biológica o clínica. Siempre reporte el tamaño del efecto junto con el valor p (Cumming, 2009; Nakagawa & Cuthill, 2007).

2.4.4 Potencia estadística

La **potencia** ($1 - \beta$) es la probabilidad de detectar un efecto real cuando existe. Es la métrica complementaria al error tipo II: si $\beta = 0{,}20$, la potencia es $1 - 0{,}20 = 0{,}80$ (80 %). Una potencia del 80 % significa que, si el efecto real existe con la magnitud esperada, el experimento lo detectará el 80 % de las veces.

La potencia depende de cuatro factores interrelacionados:

1. **Tamaño de efecto** (δ): magnitud de la diferencia real entre tratamientos. Efectos grandes son más fáciles de detectar.

2. **Nivel de significancia** (α): a mayor α, mayor potencia (pero mayor riesgo de error tipo I).

3. **Tamaño de muestra** (n): a mayor n, mayor potencia. Es el factor más controlable por el investigador.

4. **Variabilidad** (σ^2): a menor variabilidad del error, mayor potencia. Se reduce con bloqueo y estandarización.

En investigación con animales, se recomienda una potencia mínima de $1 - \beta = 0{,}80$ (80 %), lo que significa aceptar un 20 % de riesgo de no detectar un efecto real, y el tamaño de muestra necesario para alcanzar la potencia deseada se calcula con herramientas como G*Power (Cohen, 1988; Dell et al., 2002; Faul et al., 2007).

Fórmula Clave

Para un ANOVA de un factor, la potencia depende del parámetro de no centralidad

$$\lambda = \frac{r \sum_{i=1}^{a} \tau_i^2}{\sigma^2} \qquad (2.5)$$

donde r = réplicas por tratamiento, τ_i = efecto del tratamiento i, y σ^2 = varianza del error. A mayor λ, mayor potencia. El investigador puede aumentar λ incrementando r (más réplicas) o reduciendo σ^2 (mejor control experimental) (Montgomery, 2017; Ott & Longnecker, 2016).

Aplicación Veterinaria

Análisis de potencia en la práctica veterinaria. Antes de iniciar un ensayo clínico en animales, el investigador debe justificar el tamaño de muestra ante el comité de ética, justificación que requiere especificar: (1) la diferencia mínima clínicamente relevante (ej.: 3 L/día más de leche con el suplemento), (2) la variabilidad esperada (ej.: s = 5 L/día, estimada de estudios previos), (3) α = 0,05, y (4) potencia deseada = 0.80, valores con los cuales herramientas como G*Power o la ecuación de recursos del Capítulo 1 determinan n por grupo (Charan & Kantharia, 2013; Faul et al., 2007; Festing, 2018).

2.4.5 Procedimiento general de la prueba de hipótesis

Todo test de hipótesis sigue el mismo procedimiento de cinco pasos (Figura 11), independientemente de si se usa t, F , χ^2 u otro estadístico. Este procedimiento es la columna vertebral de la inferencia estadística y debe seguirse rigurosamente en cada análisis:

Concepto Clave

Los cinco pasos deben seguirse en orden. El paso 2 (fijar α) debe hacerse *antes* del paso 3 (calcular), no después. Cambiar α después de ver los resultados, por ejemplo, declarar $\alpha = 0{,}10$ porque $p = 0{,}07$, es una práctica inaceptable conocida como "mover los postes de la portería" (Wasserstein & Lazar, 2016).

2.4.6 Prueba t de Student

La prueba *t* compara las medias de **dos** grupos. Es el caso más simple de la prueba de hipótesis y un caso particular del ANOVA (cuando $a = 2$, $F = t^2$). Fue desarrollada por William Sealy Gosset, quien publicó bajo el seudónimo "Student" (Daniel & Cross, 2013; Student, 1908).

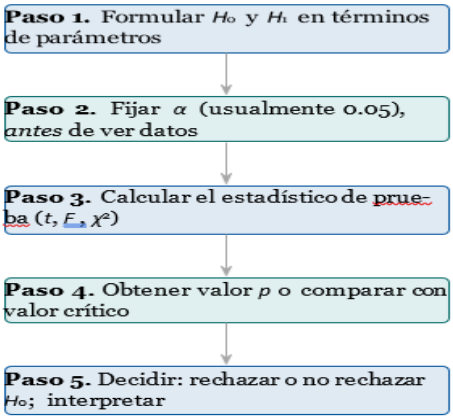


Figura 11. Procedimiento de cinco pasos para la prueba de hipótesis. El orden es estricto: los pasos 1 y 2 se definen antes de recolectar datos; los pasos 3-5 se ejecutan después.

Dato Curioso

William Sealy Gosset (1876–1937) trabajaba como químico y estadístico en la cervecería Guinness en Dublín, y como la empresa prohibía a sus empleados publicar investigaciones con su nombre real para proteger secretos industriales, recurrió al seudónimo “Student”; su prueba t, publicada en 1908 en *Biometrika*, fue revolucionaria: por primera vez se podía hacer inferencia estadística con muestras pequeñas, sin necesidad de conocer la varianza poblacional. La

distribución t converge a la normal cuando $n \to \infty$, pero con

muestras pequeñas

tiene colas más pesadas, reflejando la mayor incertidumbre.

2.4.7 Prueba t para dos muestras independientes

Fórmula Clave

Cuando se comparan dos grupos independientes con varianzas iguales (prueba *t* agrupada):

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad (2.6)$$

donde s_p es la desviación estándar agrupada (*pooled*):

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2}} \qquad (2.7)$$

Grados de libertad: $gl = n_1 + n_2 - 2$.

Supuestos:

a) Las observaciones son independientes (garantizado por aleatorización).

b) Las poblaciones son normales (verificar con Shapiro-Wilk: si $p > 0{,}05$, no se rechaza normalidad) (Shapiro & Wilk, 1965).

c) Las varianzas son iguales (verificar con Levene o Bartlett: si $p > 0{,}05$, no se rechaza homocedasticidad) (Bartlett, 1937; Levene, 1960).

Si las varianzas no son iguales, se usa la corrección de Welch, que ajusta los grados de libertad. Si la normalidad se viola severamente, se usa la prueba de Mann-Whitney como alternativa no paramétrica (Mann & Whitney, 1947).

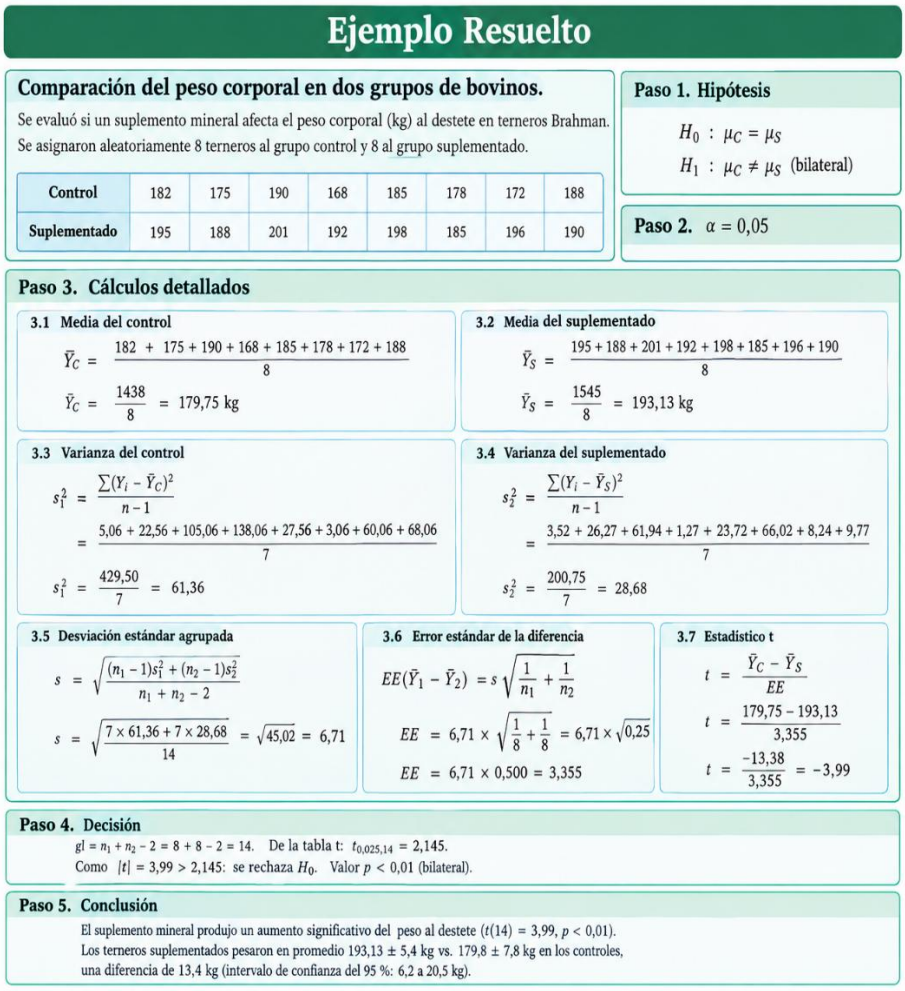


Figura 12. Ejemplo resuelto 3.

2.4.8 Prueba t para muestras pareadas

En los diseños pareados, cada unidad experimental se mide en dos condiciones (an-tes/después, izquierda/derecha, tratamiento A/tratamiento B en el mismo animal). En lugar de comparar dos grupos independientes, se trabaja con las *diferencias* individuales dentro de cada par. Esto elimina la variabilidad entre sujetos, ya que

cada animal actúa como su propio control (Zar, 2010).

Fórmula Clave

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \qquad (2.8)$$

donde $d_j = Y_{1j} - Y_{2j}$ es la diferencia para el par j, $\bar{d}$ es la media de las diferencias,
s_d es la desviación estándar de las diferencias, y gl = $n - 1$ (n = número de pares).

Hipótesis: H_0: $\mu_d = 0$ (la diferencia media poblacional es cero) vs. H_1 : $\mu_d = 0$.

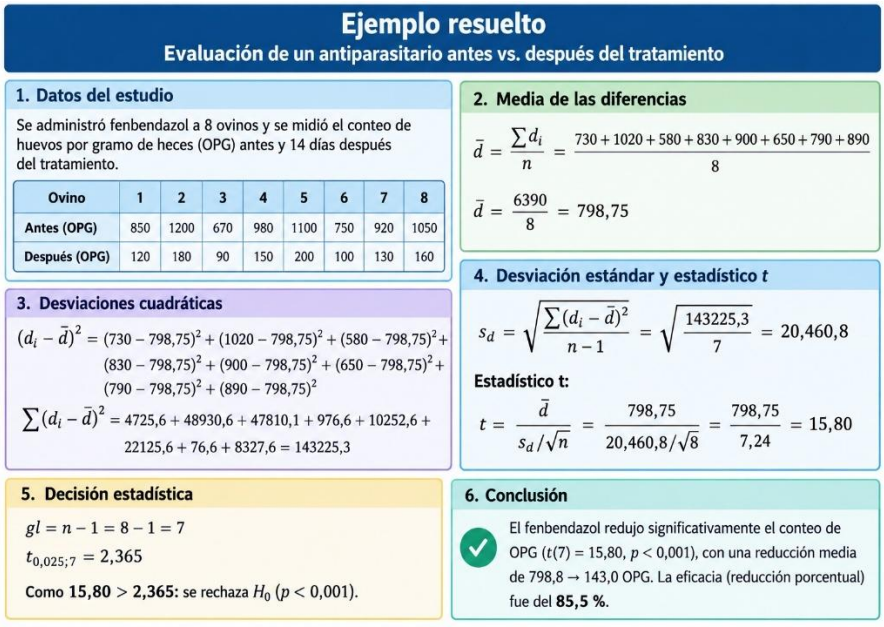


Figura 13. Ejemplo resuelto 4.

Alerta Importante

¿Prueba *t* independiente o pareada? La elección depende del diseño y no de los datos, de manera que si cada UE se mide una sola vez y los grupos son independientes → t independiente, y cuando la misma UE se mide bajo dos condiciones → t pareada; usar la prueba incorrecta produce resultados erróneos, ya que la t independiente aplicada a datos pareados ignora la correlación intra-sujeto y pierde potencia, mientras que la t pareada aplicada a datos independientes viola el supuesto de pares (Zar, 2010).

2.5 Análisis de varianza (ANOVA)

La transición de la prueba *t* al ANOVA es natural: la prueba *t* compara 2 medias; el ANOVA compara 3 o más. Pero el ANOVA no es simplemente “muchas pruebas *t* juntas”, es un marco conceptual

completamente diferente que descompone la variabilidad total en componentes, permitiendo evaluar si la variación entre tratamientos es mayor de lo esperado por azar.

2.5.1 ¿Por qué ANOVA y no múltiples pruebas t?

Si se tienen $a = 4$ tratamientos y se comparan todos los pares con pruebas t individuales, se realizan $\binom{4}{2} = 6$ comparaciones. Con $\alpha = 0{,}05$ por comparación, la probabilidad de cometer *al menos un* error tipo I es:

$$\alpha_{\text{global}} = 1 - (1 - 0{,}05)^6 = 1 - (0{,}95)^6 = 1 - 0{,}735 = 0{,}265 \quad (2.9)$$

Es decir, un 26.5 % de probabilidad de declarar al menos una diferencia significativa cuando no la hay, más de 5 veces el α individual. Esta **inflación del error tipo I** crece rápidamente con el número de tratamientos, como se detalla en la Tabla 8:

Tabla 8. Inflación del error tipo I con múltiples pruebas t (α = 0,05 por prueba).

Tratamientos (a)	**Comparaciones** $\binom{a}{2}$	α_{global}
2	1	0.050
3	3	0.143
4	6	0.265
5	10	0.401
6	15	0.537
10	45	0.901

Alerta Importante

Con 10 tratamientos ($\binom{10}{2} = 45$ comparaciones), la probabilidad de al menos un

falso positivo sube al 90 %. El ANOVA resuelve este problema realizando una sola prueba **omnibus** con α controlado al nivel declarado (0.05). La prueba t es un caso particular del ANOVA: cuando $a = 2$, $F = t^2$ y ambas pruebas dan exactamente el mismo resultado (Montgomery, 2017; Zar, 2010).

2.5.2 Lógica del ANOVA: descomposición de la variabilidad

La idea central del ANOVA es elegantemente simple: descomponer la **variabilidad total** de los datos en componentes atribuibles a diferentes fuentes, descomposición que la Tabla 9 ordena por fuente, grados de libertad y cuadrados medios. Si los tratamientos tienen un efecto real, la variabilidad *entre* tratamientos será grande en relación con la variabilidad *dentro* de los tratamientos (el error). La pregunta es: ¿cuán grande es "grande"?

Fórmula Clave

La descomposición fundamental del ANOVA:

$$SC_{Total} \quad SC_{Tr} + SC_{Err} \qquad (2.10)$$

variabilidad total entre tratamientos dentro de tratamientos

Cada observación puede descomponerse como:

$$(Y_{ij} - \bar{Y}_{..}) = (\bar{Y}_{i.} - \bar{Y}_{..}) + \qquad (2.11)$$

$$(Y_{ij} - \bar{Y}_{i.})$$

desviación total efecto del tratamiento residuo (error)

donde $\bar{Y}_{..}$ = media general, $\bar{Y}_{i.}$ = media del tratamiento *i*.

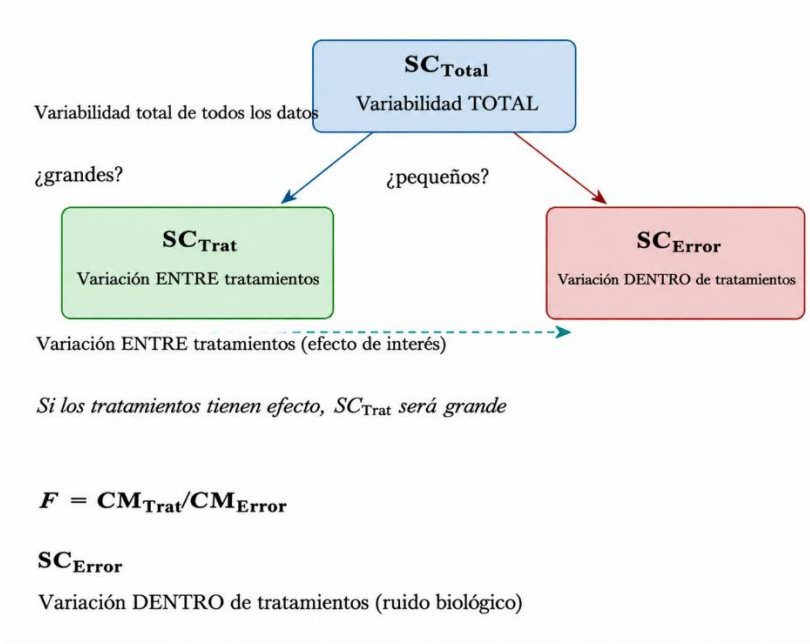


Figura 14.Descomposición de la variabilidad total en el ANOVA. La prueba F evalúa si SCTrat es "suficientemente grande" en relación con SCError, usando los cuadrados medios como estimadores de varianza.

Intuición del ANOVA: Imagine que observa 4 grupos de animales y nota que las medias difieren. ¿Es real esa diferencia? El ANOVA compara dos tipos de variabilidad: (1) la variación de las medias de grupo alrededor de la media general (SCTrat), y (2) la variación de los individuos alrededor de la media de su grupo.

Su propio grupo (SCError). Si la primera es mucho mayor que la segunda, los tratamientos probablemente tienen efecto. Si son similares, las diferencias entre medias probablemente se deben al azar (Snedecor & Cochran, 1989; Steel et al., 1997).

2.5.3 Fórmulas del ANOVA para un factor

Considérese un experimento con a tratamientos y r réplicas por tratamiento, por lo que el número total de observaciones es:

$n = a \times r$

Sea Yij la observación correspondiente al tratamiento i y a la réplica j, donde i = 1, 2, ..., a y j = 1, 2, ..., r.

Factor de corrección

El factor de corrección se calcula mediante:

$C = G^2 / n$

donde G es el gran total de todas las observaciones:

$G = \Sigma(i=1 \text{ hasta } a)\ \Sigma(j=1 \text{ hasta } r)\ Yij$

Suma de cuadrados total

La suma de cuadrados total se obtiene mediante:

$SC_Total = \Sigma(i=1 \text{ hasta } a)\ \Sigma(j=1 \text{ hasta } r)\ Yij^2 - C$

(2.13)

Suma de cuadrados de tratamientos

Para cada tratamiento se calcula primero su total:

$Ti = \Sigma(j=1 \text{ hasta } r)\ Yij$

La suma de cuadrados de tratamientos se calcula como:

$SC_Trat = (1/r)\ \Sigma(i=1 \text{ hasta } a)\ Ti^2 - C$

(2.14)

Suma de cuadrados del error

La suma de cuadrados del error se obtiene por diferencia:

$$SC_Error = SC_Total - SC_Trat$$

(2.15)

Cuadrados medios

Los cuadrados medios, que son estimadores de la varianza, se calculan mediante:

$$CM_Trat = SC_Trat / (a - 1)$$

(2.16)

$$CM_Error = SC_Error / [a(r - 1)]$$

(2.17)

Estadístico F

El estadístico F para evaluar las diferencias entre los tratamientos se calcula como:

$$F = CM_Trat / CM_Error$$

(2.18)

Si la hipótesis nula H_0 es verdadera, es decir, si todos los tratamientos tienen la misma media, CM_Trat y CM_Error son ambos estimadores de σ^2, por lo que:

$$F \approx 1$$

Si los tratamientos presentan diferencias, CM_Trat tiende a ser mayor que CM_Error y, por tanto:

$$F > 1$$

Por lo tanto, valores grandes del estadístico F proporcionan evidencia de que existen diferencias estadísticamente significativas entre las medias de los tratamientos (Montgomery, 2017).

2.5.4 Tabla ANOVA

Tabla 9. Tabla Anova.

Fuente	GL	SC	CM	F
Tratamientos	$a - 1$	SC_{trat}	$SC_{trat} / (a - 1)$	CM_{trat} / CM_{error}
Error	$a(r - 1)$	SC_{error}	$SC_{error} / [a(r - 1)]$	—
Total	$ar - 1$	SC_{total}	—	—

Concepto Clave

Verificaciones obligatorias:

1. **GL suman:** $(a - 1) + a\,(r - 1) = ar - 1$ (GL del error + GL total).
2. **SC suman:** $SC_{Trat} + SC_{Error} = SC_{Total}$.
3. $F > 0$**:** el estadístico F siempre es positivo (cociente de varianzas).

2.5.5 Distribución F y región de rechazo

El estadístico F sigue una distribución F de Fisher-Snedecor con $gl_1 = a - 1$ (numerador) y $gl_2 = a\,(r - 1)$ (denominador) grados de libertad. La distribución F es asimétrica a la derecha y siempre positiva.

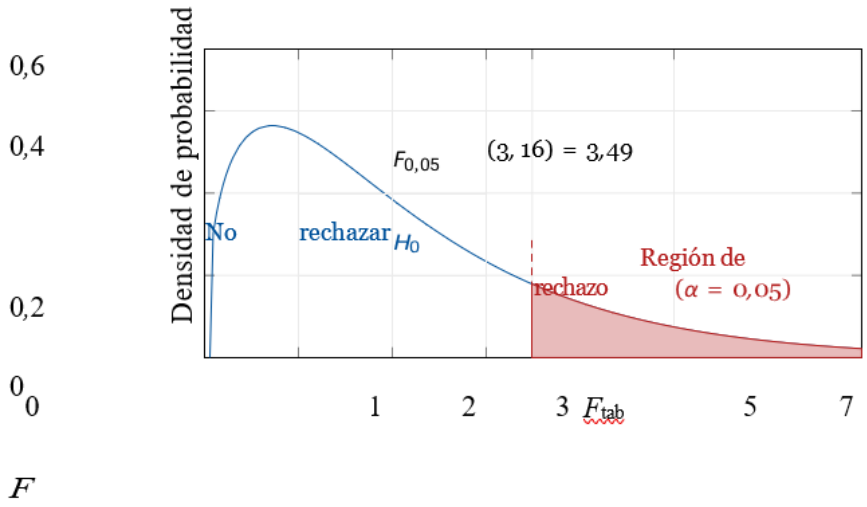


Figura 15. Distribución F con gl1 = 3 y gl2 = 16. Si Fcal > Ftab = 3,49, se rechaza H0 al nivel α = 0,05. La zona sombreada roja representa la probabilidad α de rechazar H0 cuando es verdadera.

2.5.6 Supuestos del ANOVA

El ANOVA paramétrico requiere tres supuestos sobre los datos. Estos supuestos afectan la validez de las conclusiones y deben verificarse:

- **Normalidad de los residuos:** los errores $\varepsilon_{ij} = Y_{ij} - \bar{Y}_{i.}$ se distribuyen nor-malmente. Se verifica con la prueba de Shapiro-Wilk: H_0: distribución normal. Si $p > 0{,}05$, no se rechaza normalidad (Shapiro & Wilk, 1965). El ANOVA es moderadamente robusto a desviaciones leves de normalidad, especialmente con *n* iguales.

- **Homocedasticidad (homogeneidad de varianzas):** las

1 2 a

varianzas de todos los tratamientos son iguales ($\sigma^2 = \sigma^2 = \cdots = \sigma^2$). Se verifica con la prueba de

- Levene (más robusta) o de Bartlett (más potente pero sensible a no normalidad) (Bartlett, 1937; Levene, 1960). Violaciones severas aumentan el error tipo I.

- **Independencia de las observaciones:** las observaciones de cada UE son independientes entre sí. Esto se garantiza por diseño (aleatorización correcta, identificación adecuada de la UE), no por prueba estadística. La pseudorreplicación (Capítulo 1) viola este supuesto.

Aplicación Veterinaria

Violaciones comunes en veterinaria y sus soluciones:

No normalidad por valores atípicos: Un animal enfermo o un error de medición genera un outlier. Solución: identificar la causa, excluir si justificado, o usar transformación (log, raíz).

como la variabilidad del OPG. Solución: transformación
1)) o raíz cuadrada ($\sqrt{Y}$

No independencia: Datos de animales en el mismo corral no son indepen-dientes. Solución: usar el promedio del corral como observación (Capítulo 1, pseudorreplicación).

Si los supuestos se violan severamente y no se pueden corregir, use la prueba de Kruskal-Wallis como alternativa no paramétrica (Bate & Clark, 2014; Kruskal & Wallis, 1952; Ott & Longnecker, 2016).

2.5.7 Comparaciones múltiples (post-hoc)

Si el ANOVA rechaza H_0 (*F* significativo, $p < \alpha$), sabemos que *al menos un par* de medias difiere, pero no cuáles. El ANOVA es una prueba **omnibus**: dice "hay alguna diferencia" pero no señala dónde. Las **pruebas de comparaciones múltiples** responden la pregunta "¿cuáles tratamientos difieren entre sí?", y la Tabla 10 compara las de uso más frecuente.

Prueba de Tukey (HSD)

Fórmula Clave

La prueba de Tukey (*Honestly Significant Difference*) compara todos los pares de medias controlando el error tipo I global al nivel α. La diferencia mínima significativa honesta es:

$$\text{HSD} = q_{\text{or}}(\alpha, a, \text{gl}_{\text{err}}) \times \sqrt{\frac{\text{C}_{\text{Erro}}}{\text{M}_{\text{r}}}} \quad (2.19)$$

donde q es el valor crítico de la distribución del rango estudentizado con a tratamientos y gl_{error} grados de libertad del error (Tukey, 1949).

Regla de decisión: Si $|\bar{Y}_{i.} - \bar{Y}_{j.}| \geq \text{HSD}$, los tratamientos i y j difieren significa-tivamente al nivel α.

Otras pruebas de comparaciones múltiples

Tabla 10. Comparación de pruebas de comparaciones múltiples.

Prueba	Compara	Control del error	Uso recomendado
Tukey (HSD)	Todos vs. todos	Global	General, diseños balanceados
Dunnett	Todos vs. control	Global	Comparar con testigo
Scheffé	Contrastes lineales	Global	Contrastes complejos
Bonferroni	Cualquier par	Ajuste α/k	Pocos contrastes planeados
Duncan	Todos vs. todos	Por rango	No recomendada (liberal)
LSD Fisher	Todos vs. todos	No	No recomendada (liberal)

Alerta Importante

La prueba de Duncan (1955) es la más liberal (la que más fácilmente declara diferencias significativas) y muchos estadísticos la consideran inapropiada porque no controla adecuadamente el error tipo I global. En la práctica veterinaria, se recomienda:

Tukey como prueba estándar para comparar todos los pares (la más usada). **Dunnett** cuando se compara específicamente contra un control o placebo. **Bonferroni (1936) cuando se realizan pocas comparaciones planeadas a priori** (Montgomery, 2017; Steel et al., 1997).

Letras de agrupamiento (compact letter display)

La convención de “letras de agrupamiento” facilita la comunicación de resultados, es-pecialmente en gráficos de barras. Se asignan letras a las medias ordenadas; tratamientos que comparten la misma letra no difieren significativamente.

Procedimiento:

- Ordenar las medias de menor a mayor.
- Asignar la letra “a” al grupo con la media más baja.
- Si el siguiente grupo no difiere significativamente, asignar “a” también.
- Si difiere, asignar “b”. Si no difiere del anterior pero sí del primero, asignar “ab”.
- Continuar hasta cubrir todas las medias.

2.5.8 Reporte de resultados del ANOVA

El formato estándar APA para reportar resultados de ANOVA incluye: el estadístico

F con sus grados de libertad, el valor p, y las medias con dispersión. Ejemplo:

“Se encontraron diferencias significativas entre los tratamientos (F (3, 16) =

8,42, p = 0,001). La prueba de Tukey reveló que el tratamiento A

($\bar{Y}$ = 56,2 ± 3,1 g/día) fue significativamente superior al control ($\bar{Y}$ = 48,7 ± 2,8 g/día, $p < 0{,}01$), pero no difirió del tratamiento B ($\bar{Y}$

$p = 0{,}42$).”

= 54,1 ± 3,5 g/día,

2.6 Ejemplo integrador: antiparasitarios en ovinos

Este ejemplo muestra el procedimiento completo del ANOVA, desde los datos crudos hasta la conclusión, con cada paso aritmético explícito.

2.6.1 Caso clínico

Evaluación de cuatro tratamientos antiparasitarios en ovinos

Un veterinario investigador desea evaluar la eficacia de cuatro tratamientos antiparasitarios (**T1: ivermectina, T2: albendazol, T3: levamisol y T4: control sin tratamiento**) en la reducción del conteo de huevos por gramo de heces (OPG) en ovinos **Corriedale**. Se dispone de 20 ovejas de edad y peso similar (UE homogéneas → **DCA**). Se asignan aleatoriamente 5 ovejas a cada tratamiento.

Variable respuesta: conteo de OPG a los 14 días postratamiento.

Para cumplir los supuestos de normalidad, se aplicó la transformación:

Yij∗=OPG+0.5Y_{ij}^{*}=\sqrt{OPG+0.5}

Tabla 11. Datos transformados.

Tratamiento	Y_1	Y_2	Y_3	Y_4	Y_5	T_i
T1 (Ivermectina)	3.2	2.8	3.5	2.5	3.0	**15.0**
T2 (Albendazol)	5.1	4.8	5.5	4.5	5.0	**24.9**
T3 (Levamisol)	4.2	3.8	4.5	4.0	4.3	**20.8**
T4 (Control)	8.5	9.2	8.0	9.5	8.8	**44.0**

1 Caso clínico

1. Planteamiento del problema

Evaluación de cuatro tratamientos antiparasitarios en ovinos

Un veterinario investigador desea evaluar la eficacia de 4 tratamientos antiparasitarios (T1: ivermectina, T2: albendazol, T3: levamisol, T4: control sin tratamiento) en la reducción del conteo de huevos por gramo de heces (OPG) en ovinos Corriedale. Se dispone de 20 ovejas de edad y peso similar (UE homogéneas → DCA). Se asignan aleatoriamente 5 ovejas a cada tratamiento.

Variable respuesta: conteo de OPG a los 14 días postratamiento.

Para cumplir con los supuestos de normalidad, se aplicó la transformación:

$$Y_{ij}^{\star} = \sqrt{OPG + 0{,}5}$$

2. Datos transformados

Tratamiento	Y_1	Y_2	Y_3	Y_4	Y_5	T_i
T1 (Ivermectina)	3,2	2,8	3,5	2,5	3,0	15,0
T2 (Albendazol)	5,1	4,8	5,5	4,5	5,0	24,9
T3 (Levamisol)	4,2	3,8	4,5	4,0	4,3	20,8
T4 (Control)	8,5	9,2	8,0	9,5	8,8	44,0
Total general						**104,7**

3. Hipótesis

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ (todos los tratamientos tienen el mismo efecto).

H_1: al menos un par de medias difiere.

Nivel de significancia: $\alpha = 0{,}05$.

Diseño: DCA (4 tratamientos, $r = 5$, $n = 20$).

4. Cálculos con aritmética completa

a) Medias por tratamiento

$$\bar{Y}_1 = \frac{15{,}0}{5} = 3{,}00 \qquad \bar{Y}_2 = \frac{24{,}9}{5} = 4{,}98$$

$$\bar{Y}_3 = \frac{20{,}8}{5} = 4{,}16 \qquad \bar{Y}_4 = \frac{44{,}0}{5} = 8{,}80$$

b) Gran total y media general

$$G = 15{,}0 + 24{,}9 + 20{,}8 + 44{,}0 = 104{,}7$$

$$\bar{Y} = \frac{104{,}7}{20} = 5{,}235$$

Figura 16. Caso clínico parte 1.

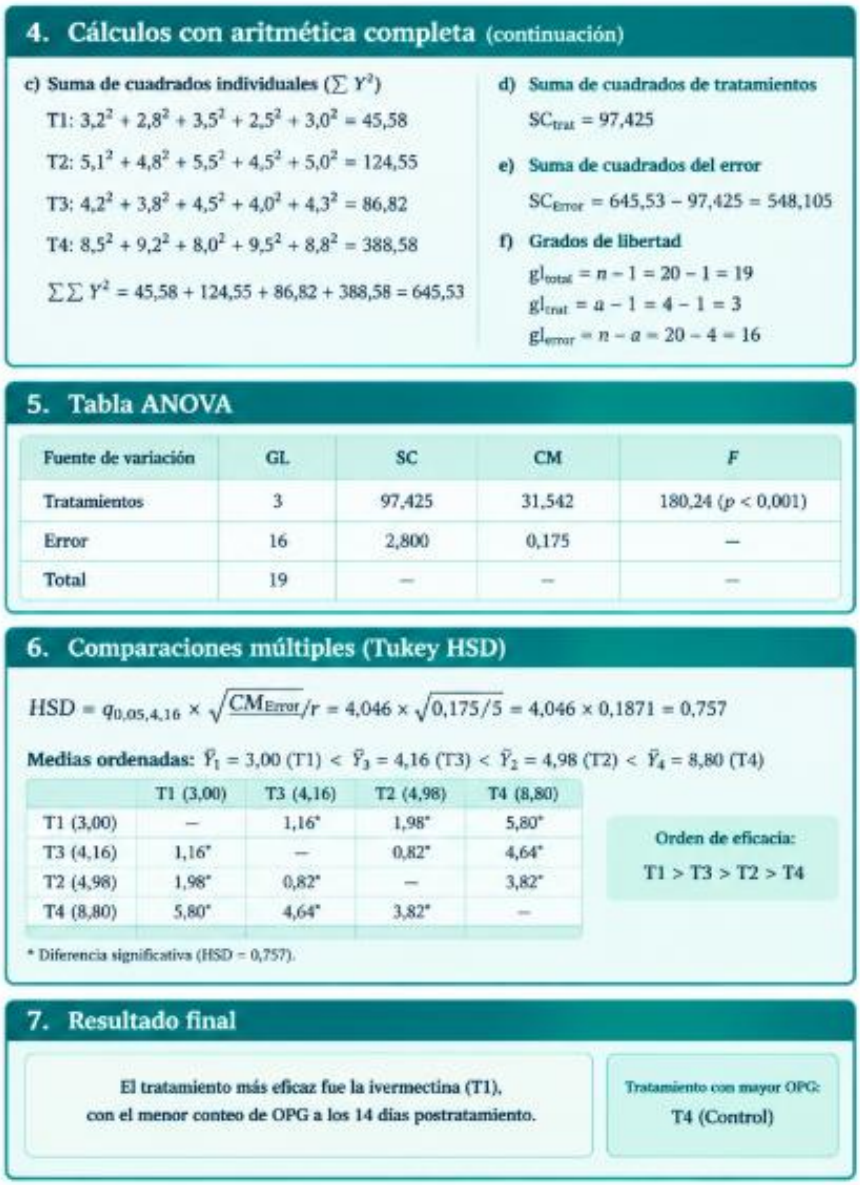


Figura 17. Caso clínico parte 2.

2.7 Tratamiento antiparasitario

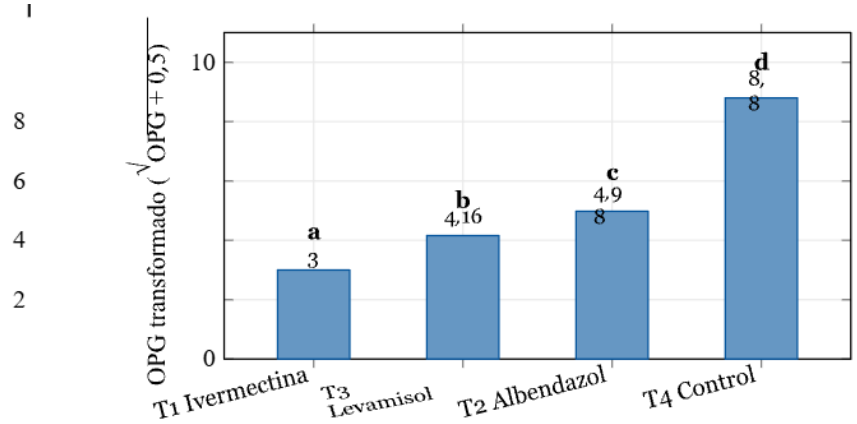


Figura 18. Medias de OPG transformados por tratamiento antiparasitario (ordenadas de menor a mayor). Letras diferentes indican diferencias significativas (Tukey HSD, p < 0,05). Barras más bajas = mayor eficacia antiparasitaria

2.7.1 Segundo ejemplo: promotores de crecimiento en pollos

Caso Clínico

Se evaluaron 3 promotores de crecimiento en pollos de engorde Ross 308 usando un DCA. La UE es el corral (no el pollo individual, ya que el promotor se mezcla en el alimento del comedero compartido). Se usaron 6 corrales por tratamiento (18 corrales total, 10 pollos por corral). Variable respuesta: ganancia diaria de peso promedio del corral (GDP, g/día) durante 21 días.

Ejemplo Resuelto

Ejemplo resuelto: ANOVA de promotores de crecimiento en pollos de engorde

1. Caso clínico / Datos de GDP (g/día) por corral

Tratamiento	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	T_i	$\bar{Y}$
T1 (Control)	42.5	44.1	41.8	43.2	42.0	44.4	258.0	43.00
T2 (Probiótico)	47.2	48.5	46.8	47.9	48.1	46.5	285.0	47.50
T3 (Enzimas)	45.0	46.2	44.5	45.8	44.8	45.7	272.0	45.33

Parámetros:

$a = 3,\ r = 6,\ n = 18$

Cálculos iniciales:

$G = 258.0 + 285.0 + 272.0 = 815.0$ $\qquad \bar{Y} = 815.0/18 = 45.28$

$C = 815.0^2/18 = 36{,}901.39$

Suma de cuadrados individuales:

T1: $42.5^2 + 44.1^2 + 41.8^2 + 43.2^2 + 42.0^2 + 44.4^2 = 11{,}099.90$

T2: $47.2^2 + 48.5^2 + 46.8^2 + 47.9^2 + 48.1^2 + 46.5^2 = 13{,}540.60$

T3: $45.0^2 + 46.2^2 + 44.5^2 + 45.8^2 + 44.8^2 + 45.7^2 = 12{,}332.86$

$\sum\sum Y^2 = 11{,}099.90 + 13{,}540.60 + 12{,}332.86 = 36{,}973.36$

Suma de cuadrados total:

$SC_{Total} = 36{,}973.36 - 36{,}901.39 = 71.97$

Suma de cuadrados de tratamientos:

$$SC_{Trat} = \frac{258^2 + 285^2 + 272^2}{6} - 36{,}901.39 = 60.78$$

Figura 19. Ejemplo Resuelto parte 1.

2. Análisis de varianza

Suma de cuadrados del error:

$SC_{Error} = 71.97 - 60.78 = 11.19$

Grados de libertad:

Tratamiento: $a - 1 = 3 - 1 = 2$
Error: $n - a = 18 - 3 = 15$
Total: $n - 1 = 17$

Tabla 12. ANOVA

Fuente de variación	GL	SC	CM	*F*	*p*
Tratamiento	2	60.78	30.39	40.74	< 0.001
Error	15	11.19	0.746	–	–
Total	17	71.97	–	–	–

Verificaciones:

GL: 2 + 15 = 17 ✓ SC: 60.78 + 11.19 = 71.97 ✓

F_{tab} (2, 15, 0.05) = 3.68. Como 40.74 > 3.68: se rechaza H_0 ($p < 0.001$).

3. Comparaciones múltiples (Tukey HSD)

$HSD = q_{0,05,3,15} \times \sqrt{CM_{Error}/r} = 3.674 \times \sqrt{0.746/6} = 3.674 \times 0.3526 = 1.295$

Medias ordenadas: T1 = 43.00 < T3 = 45.33 < T2 = 47.50

Diferencias entre tratamientos:

T1 – T3 = 2.33* T1 – T2 = 4.50* T3 – T2 = 2.17*

Nota: Todas las parejas difieren significativamente (HSD = 1.295).

Letras: $T1^a$ $T3^b$ $T2^c$

4. Conclusión

Se encontraron diferencias significativas entre los promotores de crecimiento ($F(2, 15) = 40.74$, $p < 0.001$). El probiótico (T2, 47.50 ± 0.79 g/día) produjo la mayor GDP, seguido de las enzimas (T3, 45.33 ± 0.63 g/día), ambos significativamente superiores al control (T1, 43.00 ± 1.04 g/día). Todos los tratamientos difirieron entre sí (Tukey, $p < 0.05$).

Mayor GDP:
T2 Probiótico (47.50 g/día)

Figura 20. Ejemplo Resuelto parte 2.

Aplicación Veterinaria

Nota sobre la UE en este ejemplo: obsérvese que la UE es el *corral*, no el pollo, ya que si se hubiera analizado con n = 60 pollos por tratamiento (180 pollos totales), se habría cometido pseudorreplicación: los pollos dentro del mismo corral comparten comedero y no son independientes; usar el promedio del corral como observación resulta correcto y conservador, de tal manera que con n = 6 corrales por tratamiento los GL del error son 3(6 − 1) = 15, no 3(60 − 1) = 177 (Hurlbert, 1984; Lazic, 2010).

2.8 Flujograma de decisión estadística

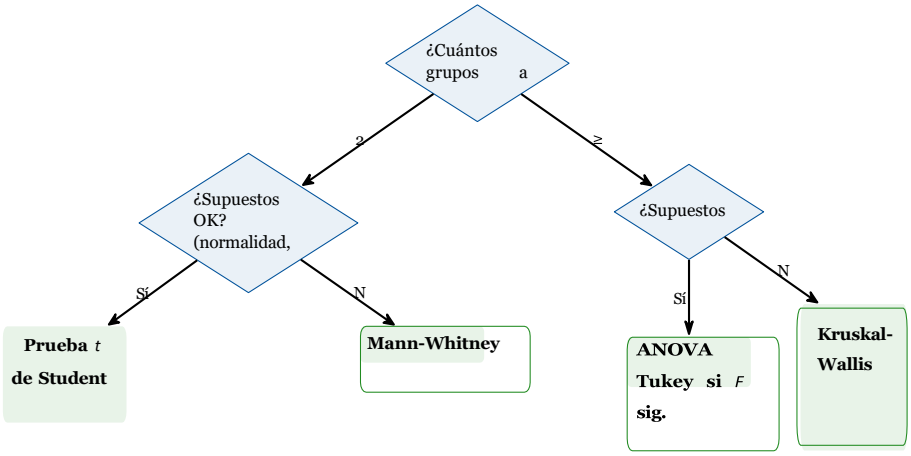


Figura 21. Flujograma de selección de prueba estadística para comparar medias de grupos independientes. Los supuestos se verifican con Shapiro-Wilk (normalidad) y Levene (homocedasticidad) (Kruskal & Wallis, 1952; Mann & Whitney, 1947).

2.9 Conceptos clave

- La **prueba de hipótesis** es el procedimiento formal de cinco pasos para decidir si los datos apoyan la existencia de un efecto.

- El **error tipo I** (α) es rechazar H_0 cuando es verdadera (falso positivo); el **error tipo II** (β) es no rechazarla cuando es falsa (falso negativo).

- El **valor** *p* no es la probabilidad de que H_0 sea verdadera; es la probabilidad de obtener los datos observados (o más extremos) si H_0 fuera cierta.

- Usar múltiples pruebas *t* infla el error tipo I; el **ANOVA** resuelve esto con una sola prueba omnibus.

- El ANOVA descompone la variabilidad total en variabilidad entre tratamientos (SC_{Trat}) y variabilidad del error (SC_{Error}). El estadístico $F = CM_{Trat}/CM_{Error}$.

- Si *F* es significativo, se usan **comparaciones múltiples** (Tukey) para identificar qué tratamientos difieren y asignar letras de agrupamiento.

- Los **supuestos** del ANOVA (normalidad, homocedasticidad, independencia) deben verificarse; las violaciones severas invalidan la prueba, y la Figura 11 ordena en un flujograma qué prueba corresponde en cada caso.

- La **potencia estadística** ($1 - \beta$) depende del tamaño de efecto, *n*, α y la variabilidad.

- La prueba *t* es un caso particular del ANOVA con $a = 2$ grupos ($F = t^2$).

- Siempre verifique: GL suman, SC suman, $F > 0$.

2.10 Actividades de consolidación

1. **Actividad 1:** Busque un artículo veterinario que reporte resultados de ANOVA. Identifique: el valor de F , los grados de libertad, el valor p, y la prueba de comparaciones múltiples utilizada. ¿Reportaron los supuestos? ¿Usaron letras de agrupamiento?

2. **Actividad 2:** Repita el cálculo completo del ejemplo integrador (antiparasitarios en ovinos), verificando cada operación aritmética. Luego cambie los datos de T4 (control) a: 6.0, 6.5, 5.8, 6.2, 5.5 y recalcule toda la tabla ANOVA y la prueba de Tukey. ¿Cambian las conclusiones?

3. **Actividad 3:** Diseñe un experimento para comparar 3 protocolos de rehidratación oral en terneros con diarrea neonatal. Defina: H_0, H_1, UE, variable respuesta, número de réplicas (justificado con la ecuación de recursos), y anticipe la estructura de la tabla ANOVA.

4. **Actividad 4:** Explique a un colega veterinario, en lenguaje no técnico, qué significa el resultado $F(3, 16) = 180{,}24$, $p < 0{,}001$ del ejemplo de antiparasitarios.

¿Qué conclusión práctica puede extraer para el manejo sanitario de sus ovinos?

2.11 Ejercicios propuestos

1. Se evaluaron 3 desinfectantes en placas con *Staphylococcus aureus*. Se usaron 6 placas por desinfectante. Los halos de inhibición (mm) fueron:

D1: 18, 22, 20, 19, 21, 17 D2: 25, 28, 24, 27, 26, 23 D3: 15, 14, 16, 13, 15, 17

Construya la tabla ANOVA completa. *Respuesta:* $SC_{Trat} = 280{,}33$, $SC_{Error} = 38{,}17$, $F = 55{,}0$, $p < 0{,}001$.

2. Con los datos del ejercicio anterior, aplique la prueba de Tukey. Use $q_{0{,}05,3,15} = 3{,}674$. *Respuesta: HSD = 2.93. Todos los pares difieren:* $D2^c > D1^b > D3^a$.

3. Se desea comparar el peso al nacimiento (kg) de terneros de dos razas: Holstein ($n = 10$: 38, 42, 35, 40, 37, 41, 39, 36, 43, 38) y Jersey ($n = 10$: 28, 32, 30, 25, 29, 31, 27, 33, 26, 30). Realice una prueba t para muestras independientes. *Respuesta:*

$\bar{Y}_H = 38{,}9$, $\bar{Y}_J = 29{,}1$, $t = 7{,}83$, $p < 0{,}001$.

4. Si se tienen 5 tratamientos, ¿cuántas comparaciones por pares se realizan? ¿Cuál sería α_{global} con pruebas t individuales ($\alpha = 0{,}05$)? *Respuesta: 10 comparaciones,* $\alpha_{global} = 0{,}401$.

5. Se evaluó el efecto de la edad al destete (21, 28, 35 días) sobre el peso a las 8 semanas en lechones ($r = 5$):

21 d: 18.2, 17.5, 19.0, 16.8, 18.5 28 d: 20.1, 21.3, 19.8, 20.5,

21.0 35 d: 22.5,

23.1, 21.8, 22.0, 23.5

Realice el ANOVA y Tukey ($q_{0,05,3,12} = 3{,}773$). *Respuesta:* $F = 46{,}5$, $p < 0{,}001$. *Todos difieren:* $35d^c > 28d^b > 21d^a$.

6. Explique por qué el ANOVA es una prueba "omnibus" y por qué un F significativo no indica cuáles tratamientos difieren.

7. ¿Cuál es la diferencia conceptual entre la prueba de Tukey y la prueba de Dunnett? Dé un ejemplo veterinario donde preferiría Dunnett.

8. Un investigador reporta $F(2, 27) = 3{,}15$, $p = 0{,}059$. ¿Qué concluye? ¿Es correcto decir "casi significativo"?

9. Describa los tres supuestos del ANOVA. ¿Cuál se verifica con Shapiro-Wilk y cuál con Levene? ¿Qué hace si se violan?

10. Un estudio con 4 tratamientos y 3 réplicas tiene $SC_T = 120$ y $SC_{Trat} = 90$. Complete la tabla ANOVA. *Respuesta:* $SC_E = 30$, $CM_{Trat} = 30$, $CM_E = 3{,}75$, $F = 8{,}0$.

11. Se midió el título de anticuerpos ($\log_2$) post-vacunación en 4 grupos de caninos con diferentes adyuvantes ($r = 4$):

A1: 5.2, 5.8, 4.9, 5.5 A2: 6.1, 6.5, 5.8, 6.3 A3: 7.2, 7.8, 6.9, 7.5 A4: 5.0, 5.3,

4.8, 5.1

Construya la tabla ANOVA. *Respuesta:* $F = 33{,}8$, $p < 0{,}001$.

12. Se pesaron 6 terneros antes y 30 días después de un tratamiento nutricional. Antes: 180, 175, 192, 168, 185, 178. Después: 195, 190, 210, 185, 200, 193. Realice una prueba *t* pareada. *Respuesta:* $\bar{d} = 15{,}33$, $t = 10{,}1$, $p < 0{,}001$.

2.12 Preguntas de autoevaluación

Selección múltiple (la respuesta correcta se indica con ←):

1. El error tipo I en la prueba de hipótesis consiste en:

a) No rechazar H_0 cuando es falsa.

b) Rechazar H_0 cuando es verdadera. ←

c) Aceptar H_1 cuando es verdadera.

d) Usar un α demasiado grande.

2. El estadístico F del ANOVA se calcula como:

a) SC_{Trat}/SC_{Error}.

b) CM_{Trat}/CM_{Error} ←

c) CM_{Error}/CM_{Trat}.

d) SC_{Total}/gl_{Total}.

3. Si el valor *p* del ANOVA es 0.03, se concluye que:

a) Todos los tratamientos son diferentes.

b) Hay un 3 % de probabilidad de que H_0 sea verdadera.

c) Existe evidencia suficiente para rechazar H_0 al nivel $\alpha = 0{,}05$. ←

d) El tratamiento más efectivo es el que tiene la media más alta.

4. La prueba de Tukey se aplica:

a) Antes del ANOVA.

b) Solo si F no es significativo.

c) Después del ANOVA, si F es significativo, para comparar pares. ←

d) Para verificar los supuestos del ANOVA.

5. La homocedasticidad se verifica con:

a) Prueba de Shapiro-Wilk.

b) Prueba de Tukey.

c) Prueba de Levene o Bartlett. ←

d) Prueba de Kruskal-Wallis.

6. Con 6 tratamientos, el número de comparaciones por pares posibles es:

a) 6.

b) 12.

c) 15. ←

d) 30.

7. Si un ANOVA tiene $a = 3$ tratamientos y $r = 8$ réplicas, los GL del error son:

a) 21. ←

b) 23.

c) 20.

d) 2.

8. Un valor de F cercano a 1 indica que:

a) Todos los tratamientos difieren significativamente.

b) La variabilidad entre tratamientos es similar a la del error. ←

c) Los supuestos están violados.

d) Se necesitan más réplicas.

9. La potencia estadística $(1 - \beta)$ aumenta cuando:

a) Disminuye el tamaño de muestra.

b) Aumenta la variabilidad del error.

c) Aumenta el tamaño de efecto o el tamaño de muestra. ←

d) Disminuye el nivel de significancia α.

10. La prueba *t* de Student para dos muestras independientes supone:

a) Normalidad e igualdad de varianzas. ←

b) Solo normalidad.

c) Solo independencia.

d) Ningún supuesto.

11. La prueba *t* pareada se usa cuando:

a) Los dos grupos son independientes.

b) Cada unidad se mide en dos condiciones (antes/después). ←

c) Hay más de dos grupos.

d) Los datos no son normales.

12. Si $F_{cal} = 2{,}50$ y $F_{tab} = 3{,}49$ ($\alpha = 0{,}05$), la decisión es:

a) Rechazar H_0.

b) No rechazar H_0. ←

c) Repetir el experimento.

d) Cambiar α a 0.10.

13. SC_{Total} se descompone en:

a) $SC_A + SC_B + SC_{AB}$.

b) $SC_{Trat} + SC_{Error}$. ←

c) $CM_{Trat} + CM_{Error}$.

d) $gl_{Trat} + gl_{Error}$.

14. El ANOVA es preferible a múltiples pruebas *t* porque:

a) Es más fácil de calcular.

b) Controla el error tipo I al nivel declarado. ←

c) No requiere supuestos.

d) Siempre da resultados significativos.

15. La diferencia entre CM_{Trat} y CM_{Error} es que:

a) CM_{Trat} estima σ^2 solo si H_0 es verdadera; CM_{Error} siempre estima σ^2. ←

b) Son iguales bajo H_1.

c) CM_{Error} siempre es mayor.

d) No hay diferencia conceptual.

Preguntas de desarrollo:

1. Explique con sus palabras por qué la inflación del error *t*ipo I es un problema cuando se realizan múltiples pruebas t y

cómo el ANOVA lo resuelve, empleando la fórmula $\alpha_{\text{global}} = 1 - (1 - \alpha)^k$ para ilustrar con un ejemplo de 5 tratamientos.

2. Un colega reporta un resultado como “no significativo (p = 0,08)” y concluye que “el tratamiento no tiene efecto.” Critique esta interpretación. ¿Qué información adicional necesitaría para evaluar correctamente el resultado? Mencione al menos tres factores.

3. Describa un escenario veterinario donde la prueba de Dunnett sea más apropia-da que la de Tukey. Formule las hipótesis y explique cómo se estructuran las comparaciones.

2.13 Glosario

ANOVA (Análisis de Varianza)

Método estadístico que compara medias de tres o más grupos descomponiendo la variabilidad total en componentes entre tratamientos y del error.

Comparaciones múltiples (post-hoc)

Pruebas que identifican cuáles pares de medias difieren después de un ANOVA significativo.

Cuadrado medio (CM)

Suma de cuadrados dividida por sus grados de libertad; es un estimador de varianza.

Descomposición de la variabilidad

Principio fundamental del ANOVA: $SC_{Total} = SC_{Trat} + SC_{Error}$.

Distribución *F*

Distribución de probabilidad del cociente de dos varianzas independientes; base del ANOVA.

Error tipo I (α)

Probabilidad de rechazar H0 cuando es verdadera; falso positivo.

Error tipo II (β)

Probabilidad de no rechazar H_0 cuando es falsa; falso negativo.

Estadístico *F*

Cociente CM_{Trat}/CM_{Error}; base de la decisión en ANOVA.

Factor de corrección (*C*)

Término G^2/n que simplifica el cálculo de las sumas de cuadrados.

Grados de libertad (GL)

Número de valores independientes que contribuyen a un cálculo estadístico.

Hipótesis alternativa (H_1)

Afirmación de que existe al menos una diferencia entre tratamientos.

Hipótesis nula (H_0)

Afirmación de que no hay diferencia entre tratamientos; es la que se somete a prueba.

Homocedasticidad

Igualdad de varianzas entre los grupos comparados; supuesto del ANOVA.

HSD (Diferencia Honestamente Significativa)

Valor umbral de la prueba de Tukey para declarar diferencia entre pares.

Inflación del error tipo I

Aumento de α_{global} al realizar múltiples pruebas sin corrección.

Letras de agrupamiento

Convención para comunicar resultados de Tukey: medias con letras iguales no difieren.

Potencia estadística (1β)

Probabilidad de detectar un efecto real cuando existe.

Prueba *t* de Student

Prueba para comparar las medias de dos grupos (independientes o pareados).

Prueba *t* pareada

Variante de la prueba *t* para datos emparejados; trabaja con diferencias individuales.

Prueba de Dunnett

Prueba post-hoc para comparar todos los tratamientos contra un control.

Prueba de Tukey (HSD)

Prueba de comparaciones múltiples que controla α_{global} al comparar todos los pares.

Prueba omnibus

Prueba que evalúa si existe alguna diferencia, sin especificar cuál (ej.: ANOVA).

Rango estudentizado (*q*)

Distribución de probabilidad usada en la prueba de Tukey; depende de a y gl_{error}.

Suma de cuadrados (SC)

Medida de variabilidad; suma de las desviaciones cuadráticas respecto a una media.

Valor *p*

Probabilidad de obtener los datos observados o más extremos si H_0 fuera verdadera.

2.14 Resumen

La **prueba de hipótesis** es el *p*rocedimien*t*o formal de cinco pasos que permite decidir si las diferencias observadas entre tratamientos son estadísticamente significativas, el cual se basa en formular una hipótesis nula (H_0: no hay diferencia) y una alternativa (H_1: existe diferencia), fijar un nivel de significancia α *antes* de recolectar datos, calcular un estadístico de prueba (t o *F*), y comparar su valor p con α. Los dos tipos de error

, tipo I (falso positivo, controlado por α) y tipo II (falso negativo, controlado por la potencia), representan los riesgos inherentes a toda decisión estadística basada en muestras (Cohen, 1988; Wasserstein & Lazar, 2016).

La **prueba *t* de Student** compara las medias de dos grupos, ya sea independientes (dos grupos de animales diferentes) o pareados (el mismo animal medido antes y después), mientras que el **análisis de varianza** (ANOVA) resulta la herramienta adecuada cuando se tienen tres o más tratamientos, ya que evita la inflación del error tipo I que ocurre al realizar múltiples pruebas t; el ANOVA descompone la variabilidad total en dos componentes fundamentales: variabilidad entre tratamientos (SC_{Trat}) y variabilidad del error (SC_{Error}), y calcula el estadístico F = CMTrat/CMError (Montgomery, 2017; Steel et al., 1997).

Si el ANOVA indica diferencias significativas ($p < \alpha$), se aplican **pruebas de com-paraciones múltiples** como Tukey (HSD) para identificar cuáles pares de tratamientos difieren, asignando letras de

agrupamiento para facilitar la comunicación de resultados (Tukey, 1949). Es esencial verificar los supuestos del ANOVA, normalidad de los residuos (Shapiro & Wilk, 1965), homocedasticidad (Levene, 1960) e independencia, y, siempre, reportar tanto la significancia estadística como la relevancia biológica de los hallazgos (Nakagawa & Cuthill, 2007).

En el siguiente capítulo se aplicarán estas herramientas dentro del marco de los **diseños experimentales clásicos**: el Diseño Completamente al Azar (DCA), el Diseño en Bloques Completos al Azar (DBCA) y el Cuadrado Latino.

CAPÍTULO III

3 DISEÑOS EXPERIMENTALES CLÁSICOS

En el capítulo anterior se desarrollaron las herramientas de inferencia estadística, la prueba de hipótesis, el ANOVA y las comparaciones múltiples, que permiten detectar diferencias entre tratamientos. Ahora se abordará la pregunta previa y fundamental:

¿cómo se organiza un experimento para que esas herramientas funcionen correctamente? Los tres diseños clásicos que se presentan en este capítulo, DCA, DBCA y Cuadrado Latino, representan las estrategias fundamentales para asignar tratamientos a unidades experimentales en investigación veterinaria.

3.1 Introducción

En la planificación de un estudio, la elección del diseño experimental es una de las decisiones más importantes y, a la vez, una de las más frecuentemente subestimadas, ya que un diseño adecuado maximiza la capacidad de detectar diferencias reales entre tratamientos, minimiza el número de unidades experimentales necesarias y garantiza la validez de las inferencias, mientras que uno inadecuado puede desperdiciar recursos, animales y tiempo sin producir conclusiones útiles (Cochran & Cox, 1957; Montgomery, 2017).

Este capítulo presenta los tres diseños experimentales más utilizados en investigación veterinaria: el Diseño Completamente al Azar (DCA), el Diseño en Bloques Completos al Azar (DBCA) y el Diseño en Cuadrado Latino (DCL). Los tres forman una

progresión natural, en la cual el DCA no controla ninguna fuente de variación extraña, el DBCA controla una y el DCL controla dos; a su vez, cada diseño se presenta con su modelo estadístico lineal, su procedimiento de aleatorización, su tabla ANOVA con fórmulas explícitas y un ejemplo veterinario completo resuelto paso a paso con toda la aritmética, a lo que se suma, al final del capítulo, una tabla comparativa y un flujograma de decisión que guían al estudiante en la selección del diseño más apropiado para cada situación (Kuehl, 2000; Mead et al., 2012; Oehlert, 2010).

La importancia práctica de este capítulo no puede exagerarse: en la literatura veterinaria, el DCA y el DBCA representan más del 80 % de los diseños publicados (Bate & Clark, 2014), de tal manera que dominarlos, junto con el ANOVA del Capítulo 2, equipa al estudiante para planificar y analizar la gran mayoría de los experimentos de su carrera profesional.

3.2 Objetivos de aprendizaje

1. Describir las características, ventajas y limitaciones de los diseños DCA, DBCA y Cuadrado Latino.

2. Escribir el modelo estadístico lineal para cada diseño e identificar cada componente.

3. Ejecutar la aleatorización correcta para cada tipo de diseño.

4. Construir la tabla ANOVA completa para DCA, DBCA y Cuadrado Latino, calculando cada celda paso a paso.

5. Calcular la eficiencia relativa del bloqueo para evaluar si la decisión de bloquear fue acertada.

6. Seleccionar el diseño experimental más adecuado según el contexto de investigación veterinaria.

7. Aplicar cada diseño a un problema experimental real, desde la planificación hasta la interpretación.

3.3 Preguntas de enfoque

1. ¿Cuándo es suficiente un DCA y cuándo se necesita bloquear?

2. ¿Qué se gana y qué se pierde al agregar bloques al diseño?

3. ¿En qué situaciones veterinarias se necesitan dos factores de bloqueo (Cuadrado Latino)?

4. ¿Cómo se evalúa si el bloqueo fue eficiente?

5. ¿Por qué el Cuadrado Latino tiene menos grados de libertad del error que el DBCA?

3.4 Diseño completamente al azar (DCA)

El Diseño Completamente al Azar es el punto de partida de todo el repertorio de diseños experimentales. Por su simplicidad y transparencia, es el diseño ideal para introducir la lógica del análisis experimental. Comprender el DCA en profundidad facilita enormemente la comprensión de los diseños más complejos que le siguen.

3.4.1 Definición y características

El **Diseño Completamente al Azar** (DCA) es el diseño más simple, en el cual los tratamientos se asignan aleatoriamente a las unidades experimentales sin ninguna restricción ni agrupación previa, y resulta apropiado cuando las UE son razonablemente homogéneas y no hay fuentes de variación extraña identificables que justifiquen el bloqueo (Cochran & Cox, 1957; Montgomery, 2017).

El DCA tiene la ventaja de la simplicidad, pues es fácil de planificar, ejecutar y analizar, y, además, al no reservar grados de libertad para bloques, maximiza los GL del error, lo que aumenta la potencia de la prueba cuando el material es homogéneo; sin embargo, su principal debilidad reside en que toda la variabilidad no explicada por los tratamientos va al error, de manera que, si las UE son heterogéneas, el error será grande y la potencia baja (Kuehl, 2000; Oehlert, 2010).

Concepto Clave

El DCA es adecuado cuando:

- Las unidades experimentales son homogéneas (misma raza, edad, peso, condición).
- No se identifica ninguna fuente de variación que justifique el bloqueo. Se trabaja en condiciones controladas (laboratorio, cámara climática). Se desea maximizar los grados de libertad del error.
- Es el diseño más común en laboratorio y el menos apropiado en campo (Kuehl, 2000; Oehlert, 2010).

Aplicación Veterinaria

Escenarios típicos para DCA en veterinaria:

- Ensayos in vitro: evaluación de desinfectantes en placas de cultivo, pruebas de susceptibilidad antibiótica, análisis de muestras de laboratorio.
- Ensayos con animales de laboratorio de cepas consanguíneas: ratones C57BL/6, ratas Wistar de la misma cepa y edad, donde la variabilidad genética es mínima.
- Evaluación de aditivos alimentarios en aves de engorde de la misma incuba-dora, misma edad, peso inicial homogéneo, y alojamiento controlado.
- Pruebas farmacológicas in vivo con animales seleccionados por peso corporal dentro de un rango estrecho (±10 %).

En todos estos casos, la homogeneidad del material experimental es la condición clave (Festing, 2006; Lazic, 2016; National Research Council, 2011).

3.4.2 Modelo estadístico lineal

Fórmula Clave

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \qquad (3.1)$$

Donde:

Y_{ij}: observación j-ésima del tratamiento i

μ: media general poblacional (constante desconocida)

τ_i: efecto del tratamiento i ($i = 1, 2, \ldots, a$); con restricción $\sum \tau_i = 0$

ε_{ij}: error experimental ($j = 1, 2, \ldots, r$)

Supuestos: $\varepsilon_{ij} \sim N(0, \sigma^2)$, independientes. Es decir, los errores son normales, con media cero, varianza constante (σ^2 igual para todos los tratamientos) e independientes entre sí (Steel et al., 1997; Winer et al., 1991).

El modelo expresa una *i*dea simple pero poderosa: cada observación es la suma de tres componentes, (1) un nivel general (μ), (2) un efecto del tratamiento (τi) y (3) una perturbación aleatoria (ε_{ij}), de tal manera que, si H_0 es verdadera (todos los tratamientos son iguales), entonces $\tau i = 0$ para todo i, y las diferencias entre observaciones se deben solo al error. La prueba *F* del ANOVA, cuya estructura de cálculo recoge la Tabla 13,

evalúa precisamente esta pregunta: ¿es la variación entre medias de tratamiento mayor de lo esperado por la variación aleatoria del error?

Alerta Importante

¿Qué pasa si los supuestos se violan? El DCA es moderadamente robusto a desviaciones de normalidad, especialmente con réplicas iguales ($r_1 = r_2 = \cdots = r$). Sin embargo, violaciones de homocedasticidad (varianzas desiguales entre tratamientos) pueden inflar el error tipo I. Soluciones: transformación de datos (logarítmica, raíz cuadrada), o uso de pruebas no paramétricas como Kruskal-Wallis (Kruskal & Wallis, 1952; Zar, 2010).

3.4.3 Aleatorización en el DCA

La aleatorización se realiza asignando cada UE a un tratamiento de forma comple-tamente aleatoria, sin restricciones:

1. Numerar todas las UE del 1 al $n = a \times r$.

2. Generar una permutación aleatoria de estos números (usando software, tabla de números aleatorios, o sorteo).

3. Asignar las primeras r UE de la permutación al tratamiento 1, las siguientes r al tratamiento 2, etc., de donde resulta un layout como el de la Figura 22.

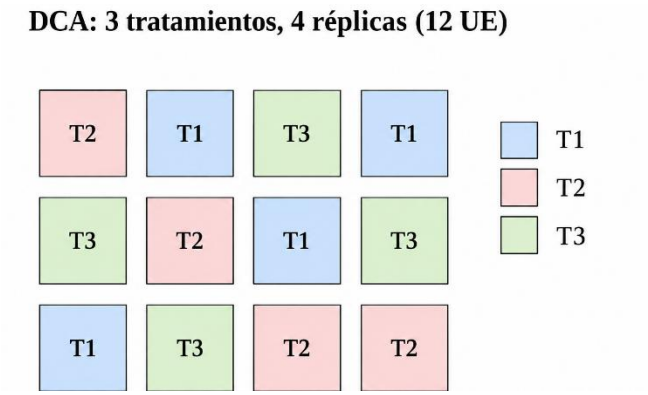


Figura 22. Layout de un DCA con 3 tratamientos y 4 réplicas. Los tratamientos se distribuyen completamente al azar entre las 12 UE. Nótese la ausencia de agrupación o restricción.g

3.4.4 Fórmulas del ANOVA para el DCA

Fórmula Clave

Con a tratamientos, r réplicas por tratamiento, $n = ar$ observaciones:

$$C = \frac{G^2}{n} \tag{3.2}$$

$$SC_{Total} = \sum_{i=1}^{a}\sum_{j=1}^{r} Y_{ij}^2 - C$$

$$SC_{Trat} = \frac{1}{r}\sum_{i=1}^{a} T_i^2 - C \tag{3.4}$$

$$SC_{Error} = SC_{Total} - SC_{Trat}$$

Donde $T_i = \sum_{j=1}^{r} Y_{ij}$ (total del tratamiento i) y $G = \sum T_i$ (gran total).

Fuente	G	SC	CM	F
Tratamiento	$a-1$	SC_{Trat}	$\frac{SC_{Trat}}{a-1}$	$\frac{CM_{Trat}}{CM_{Err}}$
Error	$a(r-1)$	SC_{Err}	$\frac{SC_{Err}}{a(r-1)}$	—
Total	$ar-1$	SC_{Total}	—	—

Figura 23. Estructura de la tabla ANOVA para el DCA.

Concepto Clave

Grados de libertad del DCA: El DCA tiene los máximos GL del error entre todos los diseños con el mismo número de UE. Con $a = 4$ y $r = 5$: $gl_{Error} = 4(5 - 1) = 16$. En un DBCA con los mismos datos: $gl_{Error} = (4 - 1)(5 - 1) = 12$ (4 GL menos). En un DCL con $a = 4$: $gl_{Error} = (4 - 1)(4 - 2) = 6$ (10 GL menos).

3.4.5 Ejemplo: evaluación de 4 antiparasitarios en caprinos

Caso clínico

1. Planteamiento del problema

Se evalúa el efecto de 4 tratamientos antiparasitarios en caprinos, que son de la misma edad (8 meses), peso similar (25 ± 2 kg) y manejo homogéneo. Se asignan aleatoriamente 6 cabras a cada tratamiento ($r = 6$, $n = 24$). Variable respuesta: reducción porcentual del conteo de OPG a los 14 días post-tratamiento.

Tratamientos:

T1: Ivermectina (0,2 mg/kg SC)

2. Datos

Tratamiento	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	T_i	$\bar{Y}_i$
T1 (Ivermectina)	95	92	97	90	93	91	558	93,00
T2 (Albendazol)	82	78	85	80	77	84	486	81,00
T3 (Levamisol)	88	85	90	87	83	89	522	87,00
T4 (Control)	12	18	8	15	20	11	84	14,00

Parámetros

$\alpha = 0,05 \quad r = 6 \quad n = 24 \quad a = 4$

Cálculos iniciales

$$G = 558 + 486 + 522 + 84 = 1650 \qquad \bar{Y} = \frac{1650}{24} = 68,75$$

$$C = \frac{1650^2}{24} = 2.722,500/24 = 113.437,50$$

Suma de cuadrados

$SC_{Total} = 138.072,00 \qquad SC_{Trat} = 36.901,39 \qquad SC_{Error} = 71.097,61$

Figura 24. Caso clinico parte 1.

3. Tabla ANOVA

Fuente de variación	GL	SC	CM	*F*	*p*
Tratamientos	3	36.901,39	12.300,46	7,81	< 0,001
Error	20	71.097,61	3.554,88	—	—
Total	23	108.000,00	—	—	—

4. Comparaciones múltiples (Tukey HSD)

$$HSD = q_{0,05,4,20} \times \sqrt{\frac{CM_{Error}}{r}} = 3,958 \times \sqrt{3.554,88/6} = 3,958 \times 24,41 = 96,58$$

Medias ordenadas:

T4 = 14,00 < T3 = 87,00 < T2 = 81,00 < T1 = 93,00

Diferencias entre tratamientos:

T1 – T2 = 12,00*	T1 – T3 = 6,00*	T1 – T4 = 79,00*
T2 – T3 = 6,00	T2 – T4 = 67,00*	T3 – T4 = 61,00*

** Diferencia significativa (HSD = 96,58).*

5. Resultado final (letras de agrupamiento)

Tukey ($p < 0,05$):

T1 (Ivermectina):	a
T3 (Levamisol):	b
T2 (Albendazol):	bc
T4 (Control):	c

Orden de eficacia:

T1 > T3 > T2 > T4

6. Conclusión

Se encontraron diferencias altamente significativas entre los tratamientos antiparasitarios ($F(3, 20) = 7,81$, $p < 0,001$). La ivermectina (T1) presentó el mayor porcentaje de reducción del OPG, seguida del levamisol (T3) y albendazol (T2), mientras que el control (T4) mostró el menor efecto. Todos los tratamientos diferenciaron significativamente entre sí (Tukey, $p < 0,05$).

Figura 25. Caso clínico parte 2.

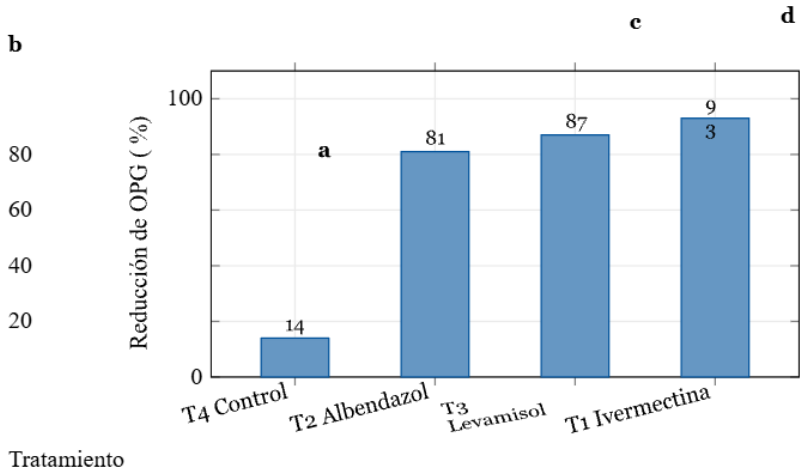


Figura 26.Reducción porcentual de OPG por tratamiento antiparasitario en caprinos Boer (DCA). Letras diferentes indican diferencias significativas (Tukey HSD, p < 0,05).

3.4.6 Ventajas y limitaciones del DCA

Tabla 12. Ventajas y limitaciones del DCA.

Ventajas	Limitaciones
Simple de planificar y ejecutar.	No controla la variación entre UE.
Flexible en número de réplicas por tratamiento.	Requiere UE homogéneas.
Máximos grados de libertad del error si hay heterogeneidad.	Error alto si hay heterogeneidad.
Análisis estadístico simple.	Potencia baja con UE heterogéneas.
Tolera datos faltantes fácilmente.	Inadecuado para ensayos de campo.

3.5 Diseño en bloques completos al azar (DBCA)

3.5.1 ¿Por qué bloquear?

Cuando las unidades experimentales no son homogéneas, la variación entre ellas incrementa el error experimental y reduce la potencia de la prueba, limitación del DCA que ya recogía la Tabla 15. El **bloqueo** agrupa las UE en bloques homogéneos, de modo que la variación entre bloques se separa del error. El resultado: un error más pequeño, una prueba más potente, y la capacidad de detectar diferencias que el DCA habría pasado por alto (Cochran & Cox, 1957; Montgomery, 2017).

Considérese un ejemplo: se desea evaluar 4 suplementos minerales en vacas lecheras que varían en su etapa de lactación (temprana, media, tardía), lo que afecta fuertemente la producción de leche, de manera que, si se usa un DCA, la variación entre etapas de lactación infla el error, mientras que al agrupar a las vacas por etapa de lactación (bloques) y asignar los 4 suplementos *dentro* de cada bloque, la variación entre etapas se "extrae" del error. El DBCA no elimina la variación entre bloques, sino que la *contabiliza* por separado, dejando un error más limpio (Kuehl, 2000; Ott & Longnecker, 2016).

Concepto Clave

El DBCA se usa cuando:

- Existe una fuente de variación **conocida** e **identificable** entre las UE. Cada bloque contiene exactamente **una UE por cada tratamiento** (bloque completo).
- La aleatorización se realiza *dentro* de cada bloque, no globalmente.
- Se espera que la variación entre bloques sea sustancial (si es mínima, el DCA es preferible).

La regla de Fisher: "bloquea lo que puedas, aleatoriza lo que no puedas controlar" (Fisher, 1935; Quinn & Keough, 2002).

Alerta Importante

¿Cuándo NO bloquear? Si las UE son genuinamente homogéneas, el bloqueo puede ser contraproducente: consume GL del error sin reducirlo. La ER será cercana a 1 (no se ganó nada) y la potencia será menor que con un DCA. Regla práctica: si duda sobre si bloquear o no, bloquee. El peor caso de un bloqueo innecesario es perder algunos GL; el peor caso de no bloquear cuando se debía es un experimento sin potencia (Lazic, 2016; Mead et al., 2012).

3.5.2 Modelo estadístico lineal

Fórmula Clave

$$Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \qquad (3.6)$$

Donde:

Y_{ij}: observación del tratamiento i en el bloque j μ: media general

τ_i: efecto del tratamiento i ($i = 1, \ldots, a$)

β_j: efecto del bloque j ($j = 1, \ldots, b$); b = número de bloques

ε_{ij}: error experimental

Nota: El modelo asume que NO hay interacción entre tratamiento y bloque. Si existiera interacción, se confundiría con el error.

Comparando con el DCA ($Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$), el DBCA agrega el término β_j. Este término absorbe la variabilidad entre bloques que en un DCA habría quedado dentro del error. El resultado es un CM_{Error} más pequeño, tal como refleja la Tabla 16 al separar la fuente de bloques.

3.5.3 Aleatorización en el DBCA

En el DBCA, la aleatorización se realiza *dentro* de cada bloque: se asignan aleato-riamente los a tratamientos a las a UE dentro del bloque.

3.5.4 Tabla ANOVA y fórmulas del DBCA

DBCA: 4 tratamientos, 5 bloques

B1	T3	T1	T4	T2
B2	T2	T4	T1	T3
B3	T1	T3	T2	T4
B4	T4	T2	T3	T1

Figura 27. Layout de un DBCA con 4 tratamientos y 5 bloques. Dentro de cada bloque, los tratamientos se asignan al azar.

Fuente de variación	GL	SC	CM	F
Bloques	$b-1$	SC_{Bloq}	$\frac{SC_{\text{Bloq}}}{b-1}$	No se prueba
Tratamientos	$a-1$	SC_{Trat}	$\frac{SC_{\text{Trat}}}{a-1}$	$\frac{CM_{\text{Trat}}}{CM_{\text{Error}}}$
Error	$(a-1)(b-1)$	SC_{Error}	$\frac{SC_{\text{Error}}}{(a-1)(b-1)}$	—
Total	$ab-1$	SC_{Total}	—	—

Figura 28. Análisis de varianza para un diseño en bloques.

Fórmula Clave

$$C = \frac{G^2}{n} \quad (n = ab) \tag{3.7}$$

$$SC_{Total} = \sum_{i=1}^{a}\sum_{j=1}^{b} Y_{ij}^2 - C \tag{3.8}$$

$$SC_{Trat} = \frac{1}{b}\sum_{i=1}^{a} T_i^2 - C \tag{3.9}$$

$$SC_{Bloq} = \frac{1}{a}\sum_{j=1}^{b} B_j^2 - C \tag{3.10}$$

$$SC_{Error} = SC_{Total} - SC_{Trat} - SC_{Bloq} \tag{3.11}$$

Donde T_i = total del tratamiento i, B_j = total del bloque j.

Eficiencia relativa del bloqueo:

$$ER = \frac{(b-1)\,CM_{Bloq} + b(a-1)CM_{Error}}{(ab-1)CM_{Error}}$$

ER > 1: el bloqueo fue eficiente. ER = 2 significa que se necesitarían el doble de UE en un DCA para la misma precisión (Casella, 2008; Montgomery, 2017).

3.5.5 Ejemplo: suplementos minerales en vacas lecheras

Caso Clínico

Se evalúan 4 suplementos minerales (T1: control, T2: selenio, T3: zinc-selenio, T4: zinc-selenio-cobre) sobre la producción de leche (L/día) en vacas Holstein. Se identifican 5 bloques según la etapa de lactación (B1: 30–60 días, B2: 61–90, B3: 91–120, B4: 121–150, B5: 151–180 días en leche).

Se analizaron 5 bloques según la etapa de lactación (B1; 30–60 días, B2: 61–90, B3: 91–120, B4: 121–150, B5: 151–180 días en leche).

Tratamiento	B1	B2	B3	B4	B5
T1 (Control)	32,5	30,1	27,8	24,2	20,5
T2 (Se)	34,2	32,5	29,6	26,8	22,3
T3 (Zn-Se)	35,8	33,1	31,2	28,5	24,1
T4 (Zn-Se-Cu)	36,5	34,8	32,0	29,1	25,0

Figura 29. Caso Clínico.

Ejemplo Resuelto

$a = 4,\ b = 5,\ n = 20.$

$C = 590{,}6^2 / 20 = 348.808{,}36 / 20 = 17.440{,}42$

$\Sigma\, Y_{ij}^2 = 32{,}5^2 + 30{,}1^2 + 27{,}8^2 + 24{,}2^2 + 20{,}5^2 + \ldots + 25{,}0^2 = 17.978{,}24$

$SC_{total} = 17.978{,}24 - 17.440{,}42 = 537{,}82$

$SC_{trat} = (135{,}1^2 + 145{,}4^2 + 152{,}7^2 + 157{,}0^2) / 5 - 17.440{,}42$

$= 17.497{,}04 - 17.440{,}42 = 56{,}62$

$SC_{bloq} = (139{,}0^2 + 130{,}5^2 + 120{,}6^2 + 108{,}6^2 + 91{,}9^2) / 4 - 17.440{,}42$

$= 17.783{,}80 - 17.440{,}42 = 343{,}38$

$SC_{error} = 537{,}82 - 56{,}62 - 343{,}38 = 137{,}82$

Verificación: $56{,}62 + 343{,}38 + 137{,}82 = 537{,}82$ □

Tabla 17. ANOVA del ensayo de suplementos minerales en vacas lecheras (DBCA)

Fuente	GL	SC	CM	F	p
Bloque (lactación)	4	343,38	85,85	—	—
Tratamiento	3	56,62	18,87	1,64	> 0,05
Error	12	137,82	11,49	—	—
Total	19	537,82			

Conclusión

No se encontraron diferencias significativas entre los suplementos ($F(3, 12) = 1{,}64$, $p > 0{,}05$), como muestra la Tabla 17, si bien el bloqueo fue altamente eficiente (ER = 2,36), ya que sin bloquear se habrían necesitado 2,36 veces más vacas para igual precisión, y la $SC_{(Bloq)} = 343{,}38$ es enorme comparad con $SC_{(Trat)} = 56{,}62$, de tal manera que la etapa de lactación explica la mayor parte de la variabilidad.

El bloqueo fue la decisión correcta aunque el tratamiento no resultó significativo (Dohoo et al., 2009; Petrie & Watson, 2013).

Figura 30. Ejercicio.

Aplicación Veterinaria

El DBCA es el diseño más utilizado en ensayos con animales de producción porque la heterogeneidad es la norma. Bloques comunes:

Nutrición: peso vivo inicial, camada, posición del corral.

Farmacología: animal (crossover), grado de infección basal.

Reproducción: número de partos, condición corporal.

Producción lechera: etapa de lactación, producción basal.

Avicultura: posición en galpón, lote de incubación.

Acuicultura: posición del estanque, flujo de agua (Bate & Clark, 2014; Thrusfield, 2018).

3.6 Diseño en cuadrado latino (DCL)

3.6.1 ¿Cuándo usar un cuadrado latino?

El DCL se emplea cuando existen **dos fuentes de variación extraña** que deben controlarse simultáneamente, y constituye una extensión del DBCA que agrega un segundo factor de bloqueo, organizado en una grilla de filas y columnas (Casella, 2008; Cochran & Cox, 1957; Mead et al., 2012).

Concepto Clave

El Cuadrado Latino controla dos fuentes de variación organizando los tratamientos en una grilla donde:

Cada tratamiento aparece exactamente **una vez en cada fila** y **una vez en cada columna**.

Número de tratamientos = número de filas = número de columnas = a.

Solo es práctico para $a = 3$ a 8 tratamientos (con $a = 3$: solo 2 GL del error).

En veterinaria, el DCL es particularmente útil en: (1) **diseños crossover**, donde filas = animales y columnas = períodos; y (2) **ensayos de campo** con gradientes en dos direcciones. El cuadrado latino fue introducido formalmente por Yates (1934) y popularizado por Fisher (1935).

3.6.2 Modelo estadístico lineal

Fórmula Clave

$$Y_{ijk} = \mu + \tau_i + \rho_j + \gamma_k + \varepsilon_{ijk} \quad (3.13)$$

Donde:

τ_i: efecto del tratamiento i ($i = 1, \ldots, a$)

ρ_j: efecto de la fila j (primer factor de bloqueo)

γ_k: efecto de la columna k (segundo factor de bloqueo)

ε_{ijk}: error experimental

Supuesto crítico: No hay interacción entre filas, columnas y tratamientos, de ahí que la Tabla 18 no incluya ninguna fuente de interacción. Si existe interacción, se confunde con el error.

3.6.3 Estructura del cuadrado latino

Cuadrado Latino 4 × 4

	Sem. 1	Sem. 2	Sem. 3	Sem. 4
Vaca 1	**A**	**B**	**C**	**D**
Vaca 2	**B**	**A**	**D**	**C**
Vaca 3	**C**	**D**	**A**	**B**
Vaca 4	**D**	**C**	**B**	**A**

Figura 31. Cuadrado Latino 4 4: 4 tratamientos de mastitis (A–D), filas = vacas, columnas = semanas. Cada tratamiento aparece exactamente una vez por fila y una vez por columna.

3.6.4 Tabla ANOVA para el DCL

Estructura de la tabla ANOVA para el Cuadrado Latino $(a \times a)$

Fuente	**GL**	**SC**	**CM**	***F***
Filas	$a-1$	$\frac{SC_{Fila}}{a-1}$	$\frac{SC_{Filas}}{a-1}$	(no se prueba)
Columnas	$a-1$	$\frac{SC_{Col}}{a-1}$	$\frac{SC_{Col}}{a-1}$	(no se prueba)
Tratamiento	$a-1$	$\frac{SC_{Trat}}{a-1}$	$\frac{SC_{Trat}}{a-1}$	$\frac{CM_{Trat}}{CM_{Error}}$
Error	$(a-1)(a-2)$	SC_{Error}	$\frac{SC_{Error}}{(a-1)(a-2)}$	
Total	a^2-1	SC_{Total}		

Figura 32. Tabla ANOVA para el DCL.

Alerta

Importante

Pocos GL del error: El DCL tiene $(a-1)(a-2)$ GL del error. Para $a = 3$: solo 2 GL (potencia muy baja). Para $a = 4$: 6 GL. Para $a = 5$: 12 GL. Esta es la principal limitación del DCL. Para cuadrados 3 × 3, se recomienda usar cuadrados latinos replicados (Cochran & Cox, 1957; Montgomery, 2017).

3.6.5 Ejemplo: tratamientos para mastitis bovina

Caso Clínico

Se evalúan 4 tratamientos para mastitis subclínica (A: cefalotporina, B: penicilina-estreptomicina, C: enllador de pasión, D: control) en vaaas Holstein. Filas = vacas (efecto genético), columnas = semanas (efecto ambiental). Variable: $\log_{10}(\text{CCS} \times 10^3/\text{mL})$.

	Semana 1	**Semana 2**	**Semana 3**	**Semana 4**
Vaca 1	A: 2.10	B: 2.35	C: 2.85	D: 3.20
Vaca 2	B: 2.50	A: 2.20	D: 3.10	C: 2.70
Vaca 3	C: 2.65	D: 3.30	A: 2.15	B: 2.45
Vaca 4	D: 3.40	C: 2.80	B: 2.55	A: 2.25

Figura 33. Caso clínico.

Ejemplo Resuelto

a = 4, $n = 16$. Totales:

$$T_A = 2{,}10 + 2{,}20 + 2{,}15 + 2{,}25 = 8{,}70, \qquad T_B = 2{,}35 + 2{,}50 + 2{,}45 + 2{,}55 = 9{,}85$$

$$T_C = 2{,}85 + 2{,}70 + 2{,}65 + 2{,}80 = 11{,}00, \qquad T_D = 3{,}20 + 3{,}10 + 3{,}30 + 3{,}40 = 13{,}00$$

$$R_1 = 10{,}50,\ R_2 = 10{,}50,\ R_3 = 10{,}55,\ R_4 = 11{,}00$$

$$K_1 = 10{,}65,\ K_2 = 10{,}65,\ K_3 = 10{,}65,\ K_4 = 10{,}60$$

$$G = 42{,}55, \qquad C = 42{,}55^2/16 = 113{,}156$$

$$\Sigma Y^2 = 115{,}808$$

$$SC_{\text{Total}} = 115{,}808 - 113{,}156 = 2{,}652$$

$$SC_{\text{Trat}} = \frac{(8{,}70^2 + 9{,}85^2 + 11{,}00^2 + 13{,}00^2)}{4} - 113{,}156 = 115{,}678 - 113{,}156 = 2{,}522$$

$$SC_{\text{Filas}} = \frac{(10{,}50^2 + 10{,}50^2 + 10{,}55^2 + 11{,}00^2)}{4} - 113{,}156 = 113{,}200 - 113{,}156 = 0{,}044$$

$$SC_{\text{Col}} = \frac{(10{,}65^2 + 10{,}65^2 + 10{,}65^2 + 10{,}60^2)}{4} - 113{,}156 = 113{,}158 - 113{,}156 = 0{,}002$$

$$SC_{\text{Error}} = 2{,}652 - 2{,}522 - 0{,}044 - 0{,}002 = 0{,}084$$

Verificación: $2{,}522 + 0{,}044 + 0{,}002 + 0{,}084 = 2{,}652$ ✓

Tabla 17. ANOVA del ensayo de mastitis (Cuadrado Latino 4 × 4)

Fuente	GL	SC	CM	*F*	*p*
Vaca (fila)	3	0.044	0.015	—	—
Semana (columna)	3	0.002	0.001	—	—
Tratamiento	3	2.522	0.841	60.07	< 0.001
Error	6	0.084	0.014	—	—
Total	15	2.652	—	—	—

$F_{\text{tab}}(3, 6, 0{,}05) = 4{,}76$. Como 60,07 >> 4,76: se rechaza H_0 ($p < 0{,}001$).

Tukey: $\text{HSD} = q_{0{,}05,4,6}\sqrt{0{,}014/4} = 4{,}896 \times 0{,}059 = 0{,}289$

Medias: $\bar{Y}_A = 2{,}175$, $\bar{Y}_B = 2{,}463$, $\bar{Y}_C = 2{,}750$, $\bar{Y}_D = 3{,}250$

$D - A = 1{,}075^*$; $D - B = 0{,}787^*$; $D - C = 0{,}500^*$;

$C - A = 0{,}575^*$;

$C - B = 0{,}287$ (NS); $B - A = 0{,}288$ (NS).

Letras: A^a B^a C^b D^c

Conclusión: Existen diferencias altamente significativas (F (3, 6) = 60,07, $p < 0{,}001$), recogidas en la Tabla 17. La cefalosporina (A) y la penicilina-estreptomicina (B) mostraron los menores CCS (no difieren entre sí), el sellador (C) fue intermedio, y el control (*D*) fue significativamente peor que todos los tratamientos activos (Bate & Clark, 2014; Snedecor & Cochran, 1989).

Figura 34. Ejemplo resuelto.

3.7 Comparación de los tres diseños

Tabla 13. Comparación de los tres diseños.

Característica	DCA	DBCA	Cuadrado Latino
Factores de bloqueo	0	1	2
Modelo	$\mu + \tau_i + \varepsilon$	$\mu + \tau_i + \beta_j + \varepsilon$	$\mu + \tau_i + \rho_i + \nu_k + \varepsilon$
GL del error (a = 4, r = 5)	16	12	6
UE necesarias	a × r	a × b	a^2
Homogeneidad requerida	—	—	Dentro de celda
Uso típico	Laboratorio	Campo, producción	Crossover, campo
Eficiencia con heterogeneidad	Baja	Alta	Alta (si 2 fuentes)
Tolerancia a datos faltantes	Alta	Moderada	Baja

Alta (global) Dentro

Aplicación Veterinaria

Reglas prácticas para seleccionar el diseño:

- **DCA:** Animales de laboratorio consanguíneos; ensayos in vitro; UE genui-namente homogéneas.
- **DBCA:** Vacas de diferente producción; cerdos de diferente camada; parcelas con gradiente; *cualquier* situación donde se identifique UNA fuente de variación relevante.
- **DCL:** Crossover farmacológico (animal × período); campo con dos gradientes; evaluación de dietas en caballos (caballo × semana).

En la duda: elija DBCA antes que DCA, es más seguro (Dohoo et al., 2009; Quinn & Keough, 2002; Thrusfield, 2018; Winer et al., 1991).

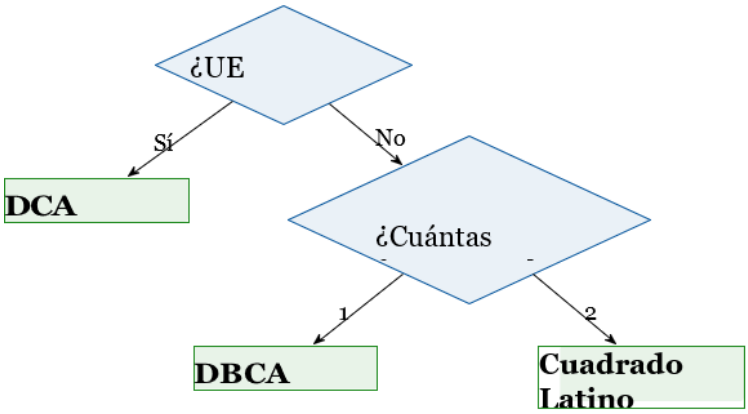


Figura 35. Flujograma de selección entre DCA, DBCA y Cuadrado Latino.

Concepto Clave

Resumen de GL del error por diseño (con a = 4 tratamientos): DCA ($r = 5$): $gl_E = 4(5 - 1) = 16$

DBCA ($b = 5$): $gl_E = (4 - 1)(5 - 1) = 12$

DCL: $gl_E = (4 - 1)(4 - 2) = 6$

El DCA tiene más GL, pero si el material es heterogéneo, esos GL contienen variación inútil. El DBCA y DCL sacrifican GL para reducir CM_{Error}, un intercambio que vale la pena cuando SC_{Bloq} es grande (Box et al., 2005; Lazic, 2016; Zar, 2010).

3.8 Conceptos clave

- El **DCA** asigna tratamientos sin restricción. Adecuado solo con UE homogéneas. Máximos GL del error.

- El **DBCA** introduce un factor de bloqueo para controlar una

fuente de variación conocida, reduciendo el error.

- El **Cuadrado Latino** controla dos fuentes de variación, pero requiere *a* = filas= columnas = tratamientos.
- La **eficiencia relativa** cuantifica la ganancia del bloqueo: ER > 1 indica que bloquear fue útil.
- El bloque NO se prueba con *F* : es un factor de control, no de interés.
- Cada diseño tiene GL del error distintos: DCA > DBCA > DCL.
- Más GL del error no siempre es mejor: lo que importa es que CM_{Error} sea pequeño.
- La decisión de bloquear se toma *antes* del experimento, basada en conocimiento previo.

3.9 Actividades de consolidación

1. **Actividad 1:** Diseñe un DBCA para evaluar 3 niveles de proteína en cerdos en crecimiento. Defina el factor de bloqueo, justifique su elección, dibuje el layout de aleatorización con 4 bloques.

2. **Actividad 2:** Un colega realizó un ensayo con 4 herbicidas en 5 parcelas por herbicida (DCA), en el cual los CV fueron altos (> 30 %). Proponga un rediseño con DBCA, identifique el factor de bloqueo y explique por qué la ER probablemente

será alta.

3. **Actividad 3:** Construya un cuadrado latino 3 × 3 para evaluar 3 protocolos de sincronización de celo en vacas, bloqueando por vaca y por mes. Verifique la propiedad del cuadrado latino.

4. **Actividad 4:** Compare los GL del error de un DCA, DBCA y DCL para a = 5 tratamientos. ¿Cuál tiene más GL? ¿Cuál tendría menor error si el material es heterogéneo en una dimensión?

3.10 Ejercicios propuestos

1. Se evaluaron 4 niveles de vitamina E (0, 100, 200, 400 UI/kg) en trucha ar-coíris (DCA, $r = 5$ estanques). Ganancias (g): 0: 120,115,125,118,122; 100: 135,128,140,132,138; 200: 142,148,138,145,140; 400: 140,135,145,138,142. Realice ANOVA y Tukey ($q_{0,05,4,16} = 4{,}046$). *Resp.:* $F = 18{,}3$, $p < 0{,}001$.

2. Con 3 antibióticos y 4 bloques (peso): A1: 3.2,3.5,3.8,4.0; A2: 2.8,3.0,3.2,3.5; A3: 2.5,2.8,3.0,3.2 ($\log_{10}$ UFC/mL). ANOVA del DBCA y eficiencia relativa del bloqueo.

3. GL del error de un cuadrado latino 5 × 5 y uno 3 × 3. ¿Cuál tiene más potencia? *Resp.: 12 y 2 respectivamente.*

4. DCL 4 × 4: SCT = 180, SCDieta = 85, SCVaca = 50, SCPeríodo = 20. Complete la tabla ANOVA. Resp.: SCE = 25, F = 6,80.

5. Diferencia en GL del error entre DCA (a = 4, r = 5), DBCA (a = 4, b = 5) y DCL (a = 4). ¿Siempre es mejor más GL?

6. ¿Cuándo DBCA en vez de DCA? (a) ratones consanguíneos; (b) vacas de diferente paridad; (c) pollitos de igual incubadora; (d) parcelas en colina. Resp.: (b) y (d).

7. Si ER ≈ 1 en un DBCA, ¿qué significa? ¿Se pierde potencia?

8. DCL 4 × 4: si un animal se retira en semana 3, ¿qué problemas genera?

9. DBCA para 5 vacunas en perros con 4 bloques: ¿cuántos perros? Resp.: 20.

10. ¿Por qué el DCL asume no interacción fila-columna-tratamiento?

11. Se evaluaron 3 desparasitantes en caprinos (DCA, r = 7): D1: 92,88,95,90,87,93,91; D2: 78,82,75,80,85,77,83; D3: 65,70,62,68,72,60,66. ANOVA completa y Tukey ($q_{0,05,3,18} = 3,609$).

12. DCA con a = 4 y r = 3: GL del error = 8. ¿Suficiente? Proponga alternativa. Resp.: Insuficiente (E < 10). Aumentar r o usar crossover.

3.11 Preguntas de autoevaluación

Selección múltiple (← = correcta):

1. DCA es apropiado cuando: (a) UE heterogéneas; (b) dos fuentes de variación; (c) UE homogéneas ←; (d) campo.

2. Propósito del bloque en DBCA: (a) tratamiento adicional; (b) reducir error controlando variación ←; (c) aumentar GL; (d) reemplazar aleatorización.

3. GL del error en DCL 5 × 5: (a) 20; (b) 16; (c) 12 ←; (d) 24.

4. ER > 1 indica: (a) bloqueo ineficiente; (b) ER = 0; (c) bloqueo eficiente ; (d) ER = 1.

5. Bloque NO se prueba con *F* porque: (a) sin GL; (b) factor de control ; (c) siempre significativo; (d) sin varianza.

6. En $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$, τ_i es: (a) error; (b) bloque; (c) tratamiento ←; (d) media.

7. Diseño con a = filas = columnas = tratamientos: (a) DCA; (b) DBCA; (c) DCL

←; (d) Factorial.

8. DCA con a = 4, r = 6: GL error: (a) 20 ←; (b) 24; (c) 23; (d) 3.

9. DBCA tiene menos GL que DCA porque: (a) más tratamientos; (b) GL del bloque se restan ←; (c) error menor; (d) aleatorización diferente.

10. β_j en DBCA es: (a) tratamiento; (b) error; (c) bloque ←; (d)

interacción.

11. ER = 3.5 significa: (a) ineficiente; (b) 3.5 veces más UE en DCA; (c) 3.5 bloques; (d) 3.5 veces más efecto.

12. En DCL, cada tratamiento aparece: (a) múltiples veces/fila; (b) una vez/fila y una vez/columna ←; (c) solo diagonal; (d) posiciones fijas.

13. Supuesto crítico del DCL: (a) > 5 tratamientos; (b) no interacción fila-col-trat

; (c) filas = columnas; (d) error = 0.

14. DCL 3 × 3 requiere: (a) 3 UE; (b) 6; (c) 9 ←; (d) 12.

15. DBCA a = 5, b = 4: GL error: (a) 12 ←; (b) 15; (c) 19; (d)

Preguntas de desarrollo:

1. Compare DCA y DBCA para evaluar 4 aditivos en pollos. ¿Qué factor de bloqueo? Escriba ambos modelos y compare GL del error.

2. DCL 3 x 3: F = 2,5 con GL(2, 2) y Ftab = 19,0. ¿Conclusión? ¿Problema del diseño? ¿Recomendación?

3. Describa un escenario veterinario donde el DCL sea óptimo. Justifique filas y columnas.

3.12 Glosario

Aleatorización dentro de bloques

Asignación aleatoria de tratamientos a UE dentro de cada bloque en un DBCA.

Aleatorización irrestricta

Asignación aleatoria sin restricciones; característica del DCA.

Bloque

Grupo de UE homogéneas que contiene todos los tratamientos.

Crossover

Diseño donde cada sujeto recibe todos los tratamientos secuencialmente; caso especial del DCL.

Cuadrado Latino (DCL)

Diseño con dos factores de bloqueo; cada tratamiento aparece una vez por fila y columna.

DCA

Diseño Completamente al Azar; sin bloqueo, asignación totalmente aleatoria.

DBCA

Diseño en Bloques Completos al Azar; un factor de bloqueo.

Diseño balanceado

Igual número de réplicas por tratamiento.

Diseño desbalanceado

Diferente número de réplicas; complica el análisis.

Eficiencia relativa (ER)

Cuánto ganó el bloqueo respecto al DCA; ER > 1 = eficiente.

Factor de bloqueo

Variable que genera heterogeneidad y se controla por agrupación.

Factor de corrección (*C*)

G^2/n; simplifica el cálculo de sumas de cuadrados.

GL del error

Información disponible para estimar variabilidad residual; varía por diseño.

Homocedasticidad dentro de bloques

Varianzas iguales entre tratamientos dentro de cada bloque.

Interacción tratamiento-bloque

Viola supuesto de aditividad del DBCA/DCL.

Layout experimental

Disposición física de tratamientos en las UE tras aleatorización.

Modelo estadístico lineal

Descomposición de cada observación en efectos.

Cuadrado latino replicado

Repetición del CL completo para aumentar GL del error.

Suma de cuadrados de bloques

Variabilidad atribuida a diferencias entre bloques.

Restricción del CL

a = filas = columnas = tratamientos.

Varianza del error (CM_{Error})

Denominador del estadístico F; estimador del error.

Período de lavado

Tiempo entre tratamientos en crossover.

Potencia del diseño

Capacidad de detectar diferencias; depende de GL y CM_{Error}.

SC de tratamiento

Variabilidad atribuida a diferencias entre medias de tratamiento.

3.13 Resumen

Los tres diseños clásicos (**DCA**, **DBCA** y **Cuadrado Latino**) form*a*n una progresión en el control de la variabilidad experimental, en la cual el DCA es el más simple, pues asigna tratamientos sin restricciones y resulta adecuado cuando las UE son homogéneas, a la vez que maximiza los GL del error, aunque no controla fuentes de variación extrañas. Con el DBCA se introduce un factor de bloqueo que separa una fuente de variación conocida del error y aumenta la potencia, ganancia que la eficiencia relativa cuantifica, mientras que el Cuadrado Latino controla dos fuentes simultáneamente a costa de menos GL del error y de la restricción a = filas = columnas (Cochran & Cox, 1957; Montgomery, 2017).

La selección depende de la estructura del material experimental: homogéneo → DCA; una fuente de variación → DBCA; dos fuentes → DCL, sin que pierda vigencia la regla de Fisher, "bloquea lo que puedas, aleatoriza lo que no puedas" (Box et al., 2005; Fisher, 1935; Lazic, 2016; Yates, 1934; Zar, 2010).

En el capítulo siguiente se abordarán las factoriales (dos o más factores simultá-neamente, con detección de interacciones) y la **parcela dividida** (factores con distinta facilidad de aleatorización).

CAPÍTULO IV

4 EXPERIMENTOS FACTORIALES

Los capítulos anteriores presentaron diseños con un solo factor de interés: un grupo de tratamientos que se evalúan contra un error experimental. Sin embargo, muchas preguntas de investigación veterinaria involucran dos o más factores simultáneamente.

¿Afecta el nivel de proteína en la dieta la ganancia de peso? ¿Y el nivel de energía?

¿Y la combinación de ambos? Los experimentos factoriales permiten responder estas preguntas con eficiencia y, además, detectar un fenómeno que los diseños unifactoriales no pueden revelar: la interacción entre factores.

4.1 Introducción

Un **experimento factorial** es aquel en el que se investigan simultáneamente dos o más factores, evaluando todas las combinaciones posibles de sus niveles, lo que hace de los diseños factoriales una opción más eficiente que los estudios de un solo factor a la vez, porque cada observación contribuye a estimar los efectos de *todos* los factores y sus interacciones. En este contexto, el capítulo desarrolla los conceptos de efecto principal e interacción, presenta el modelo y la tabla ANOVA del factorial completa, introduce el diseño en parcela dividida para factores con distinta facilidad de aleatorización y ofrece una visión panorámica de diseños avanzados que el estudiante debe reconocer aunque excedan el alcance cuantitativo de este texto

(Box et al., 2005; Montgomery, 2017).

La capacidad de detectar **interacciones** es la ventaja más importante del diseño factorial, entendiendo por interacción la situación en que el efecto de un factor depende del nivel del otro. En veterinaria, las interacciones son biológicamente comunes y clínicamente relevantes: un antibiótico puede ser eficaz a dosis baja en una especie pero no en otra, del mismo modo que un suplemento nutricional puede mejorar la producción solo en animales con cierta base genética, de tal manera que ignorar las interacciones, como ocurre cuando se evalúa un factor a la vez, puede llevar a conclusiones incompletas o incorrectas (Kuehl, 2000; Ott & Longnecker, 2016).

4.2 Objetivos de aprendizaje

1. Definir qué es un experimento factorial y diferenciarlo de un diseño unifactorial.

2. Distinguir entre efecto principal e interacción e interpretar gráficos de interacción.

3. Escribir el modelo estadístico y construir la tabla ANOVA para una factorial *b* en DCA.

4. Calcular paso a paso cada suma de cuadrados del factorial, incluyendo la interac-ción.

5. Identificar cuándo se requiere un diseño en parcela dividida y describir su estructura de doble error.

6. Seleccionar el diseño experimental más apropiado entre DCA, DBCA, Cuadrado Latino, factorial y parcela dividida según el contexto.

7. Reconocer conceptualmente los diseños de medidas repetidas, crossover, aumenta-dos y anidados.

4.3 Preguntas de enfoque

1. ¿Por qué es más eficiente evaluar dos factores juntos en un factorial que en dos experimentos separados?

2. ¿Qué significa que “hay interacción” entre dos factores y por qué cambia la interpretación de los efectos principales?

3. ¿En qué situaciones un factor no puede aleatorizarse a nivel de la unidad experi-mental pequeña y cómo lo resuelve la parcela dividida?

4. ¿Cómo se decide si se necesita un diseño simple o uno avanzado?

4.4 Experimentos factoriales: conceptos básicos

4.4.1 ¿Qué es un experimento factorial?

Concepto Clave

Un **experimento factorial** evalúa simultáneamente dos o más factores, inclu-yendo *todas las combinaciones* de sus niveles como tratamientos. En una factorial

$a \times b$, con a niveles del factor A y b niveles del factor B, hay a

$\times b$ tratamientos
(combinaciones). Cada observación contribuye a estimar los efectos de A, de B y de la interacción $A \times B$ (Cochran & Cox, 1957; Montgomery, 2017).

La diferencia fundamental entre un factorial y un diseño unifactorial es que en el factorial, los "tratamientos" no son niveles de un solo factor sino *combinaciones* de niveles de dos o más factores. Por ejemplo, en un factorial 2 × 3 (2 antibióticos × 3 dosis), los 6 tratamientos son: (Antibiótico 1, Dosis baja), (Antibiótico 1, Dosis media), (Antibiótico 1, Dosis alta), (Antibiótico 2, Dosis baja), (Antibiótico 2, Dosis media), (Antibiótico 2, Dosis alta).

Aplicación Veterinaria

Un nutricionista avícola se propone evaluar el efecto de 2 niveles de proteína (18 % y 22 %) y 3 niveles de energía metabolizable (2800, 3000, 3200 kcal/kg) sobre la ganancia de peso en pollos, de tal manera que, bajo un enfoque unifactorial, necesitaría dos experimentos

separados. En un factorial 2×3, evalúa las 6 combinaciones en un solo experimento,
con menos animales y más información: obtiene los efectos de la proteína, los efectos de la energía, *y* la interacción proteína × energía, todo con el mismo

conjunto de datos (Quinn & Keough, 2002).

4.4.2 Terminología factorial

Factor: Variable independiente con dos o más niveles (ej.: tipo de antibiótico, dosis).

Nivel: Cada valor específico de un factor (ej.: dosis baja, media, alta).

Combinación de tratamientos:

Cada par (o trío, etc.) específico de niveles de los factores. En un 2 × 3: (A_1B_1), (A_1B_2), (A_1B_3), (A_2B_1), (A_2B_2), (A_2B_3), combinaciones cuyo número crece deprisa al añadir factores y niveles, según cuantifica la Tabla 20.

Efecto principal:

Efecto promedio de un factor, promediando sobre todos los niveles del otro factor.

Interacción: Cuando el efecto de un factor depende del nivel del otro factor.

Tabla 14. Número de combinaciones de tratamientos en diferentes factoriales.

Factorial	Factores	Niveles	Combinaciones
2 × 2	2	2, 2	4
2 × 3	2	2, 3	6
3 × 3	2	3, 3	9
2 × 2 × 2	3	2, 2, 2	8
3 × 4	2	3, 4	12

4.4.3 Efecto principal e interacción

Fórmula Clave

El **efecto principal** de un factor es la diferencia entre las medias marginales de sus niveles.

La **interacción** $A \times B$ mide la desviación de los efectos combinados respecto a lo esperado por la suma de los efectos principales individuales. Formalmente:

$$(\alpha\beta)_{ij} = \mu_{ij} - \mu - \alpha_i - \beta_j \quad (4.1)$$

Si $(\alpha\beta)_{ij} = 0$ para todo i, j, no hay interacción; los efectos son **aditivos** (Kuehl, 2000; Montgomery, 2017).

La interacción es el concepto más difícil y más importante de este capítulo. Para comprenderla intuitivamente: si el efecto de la proteína sobre la ganancia de peso es el mismo con energía baja que con energía alta, no hay interacción. Pero si la proteína alta solo mejora la ganancia cuando la energía también es alta, hay interacción, el efecto de la proteína *depende* del nivel de energía, contraste que la Figura 37 y 37 traduce en líneas paralelas o cruzadas.

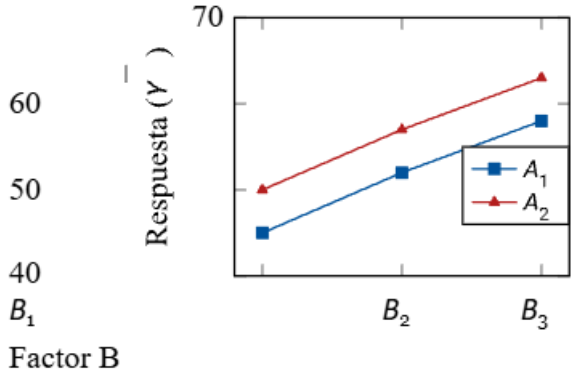


Figura 36. (a) Sin interacción

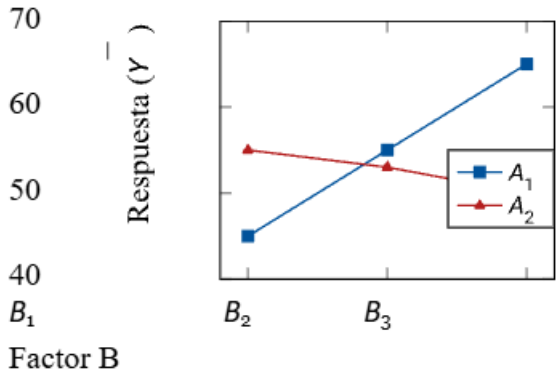


Figura 37. (b) Con interacción.

Nota. Gráficos de interacción. (a) Líneas paralelas indican ausencia de interacción: el efecto de B es consistente en ambos niveles de A. (b) Líneas que se cruzan indican interacción: el efecto de B depende del nivel de A.

Alerta Importante

Cuando la interacción es significativa, los **efectos principales** pierden significado práctico y pueden ser engañosos: si el antibiótico A es mejor que B con la dosis baja, pero peor con la dosis alta, decir "en promedio A es mejor" no tiene sentido biológico, de tal manera que, en presencia de interacción significativa, se deben interpretar los **efectos simples** (el efecto de un factor en cada nivel fijo del otro) (Montgomery, 2017; Ott & Longnecker, 2016).

Regla práctica: Siempre evalúe la interacción *primero*. Si es significativa, analice efectos simples. Si no es significativa, interprete los efectos principales directa-mente.

4.4.4 Ventajas del diseño factorial

a) **Eficiencia:** cada réplica contribuye a estimar todos los efectos. Una factorial 2 3 con r = 4 usa 24 UE para estimar 2 efectos principales + 1 interacción; dos experimentos unifactoriales separados necesitarían más UE para la misma potencia.

b) **Detección de interacciones:** imposible con diseños unifactoriales.

c) **Alcance de inferencia:** las conclusiones son válidas sobre el rango de combina-ciones evaluadas, no solo para niveles individuales.

d) **Economía:** requiere menos UE que dos experimentos separados para la misma potencia (Box et al., 2005; Casella, 2008).

Concepto Clave

La **eficiencia oculta** de la factorial: en un factorial 2 × 2 con *r* réplicas por

celda, cada efecto principal se estima con 2r observaciones por nivel (todas las observaciones de A_1, independientemente de B), lo cual equivale a un experimento unifactorial de A con 2r réplicas por nivel, aunque sin costo adicional, porque las mismas observaciones también estiman B y la interacción (Montgomery, 2017).

4.5 Modelo y ANOVA de la factorial a × b en DCA

4.5.1 Modelo estadístico lineal

Fórmula Clave

Para un factorial $a \times b$ con r réplicas en un DCA:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \quad (4.2$$

Donde:

μ: media general

α_i: efecto del nivel i del factor A ($i = 1, \ldots, a$) β_j: efecto del nivel j del factor B ($j = 1, \ldots, b$) $(\alpha\beta)_{ij}$: efecto de la interacción $A \times B$

ε_{ijk}: error experimental ($k = 1, \ldots, r$)

$n = abr$ observaciones totales

Supuestos: $\varepsilon_{ijk} \sim N(0, \sigma^2)$, independientes (Steel et al., 1997; Winer et al., 1991).

Comparando con el modelo del DCA unifactorial ($Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$), el modelo factorial reemplaza el efecto de tratamiento τ_i por tres componentes: α_i (efecto de A), β_j (efecto de B) y $(\alpha\beta)_{ij}$ (interacción). Esto permite descomponer la SC_{Trat} en tres partes, cada una probada independientemente con su propio estadístico F, tal como se organiza en la Tabla 21.

4.5.2 Tabla ANOVA de la factorial

Tabla 15. Estructura de la tabla ANOVA para un factorial a × b en DC.

Fuente	GL	SC	CM	F
Factor A	$a - 1$	SC_a	$SC_a / (a - 1)$	CM_a / CM_e
Factor B	$b - 1$	SC_β	$SC_\beta / (b - 1)$	CM_β / CM_e
A × B	$(a - 1)(b - 1)$	$SC_{a\beta}$	$SC_{a\beta} / [(a - 1)(b - 1)]$	$CM_{a\beta} / CM_e$
Error	$ab(r - 1)$	SC_e	$SC_e / [ab(r - 1)]$	—
Total	$abr - 1$	SC_t	—	—

Concepto Clave

Verificación de GL en el factorial: Los GL deben sumar: $(a - 1) + (b - 1) + (a - 1)(b - 1) + ab(r - 1) = abr - 1$. Para un factorial 2×3 con $r = 4$: $1 + 2 + 2 + 18 = 23 = 24 - 1$ ✓. Esta verificación es una herramienta útil para detectar errores de cálculo.

4.5.3 Fórmulas de sumas de cuadrados

Fórmula

$$C = \frac{G^2}{abr} \tag{4.3}$$

$$SC_T = \sum_i \sum_j \sum_k Y_{ijk}^2 - C \tag{4.4}$$

$$SC_A = \frac{1}{br} \sum_i A_i^2 - C \tag{4.5}$$

$$SC_B = \frac{1}{ar} \sum_j B_j^2 - C \tag{4.6}$$

$$S_{AB(\text{subtota}} = \frac{1}{r} \sum_i \sum_j (AB_{ij}^2 - C \tag{4.7}$$

$$SC_{AB} = SC_{AB(\text{subtotales})} - \tag{4.8}$$

$$SC_E = SC_T - SC_A - \tag{4.9}$$

Donde A_i = total del nivel i del factor A, B_j = total del nivel j de B, $(AB)_{ij}$ = total de la celda (i, j), G = gran total.

4.5.4 Ejemplo: factorial 2 × 3 en nutrición porcina

Caso Clínico

Se evalúa el efecto de 2 niveles de proteína (A1: 16 %, A2: 20 %) y 3 niveles de lisina. (B1: 0.8 %, B2: 1.0 %, B3: 1.2 %) sobre la ganancia diaria de peso (GDP) en cerdos en crecimiento. Diseño: factorial 2 × 3 en DCA con r = (cerdos) por combinación. Total: 2 × 3 × 4 = 24.

Proteína	Lisina	Y_1	Y_2	Y_3	Y_4	$(AB)_{ij}$
A1 (16 %)	B1 (0.8 %)	620	635	610	625	2490
	B2 (1.0 %)	680	670	695	685	2730
	B3 (1.2 %)	700	715	690	705	2810
A2 (20 %)	B1 (0.8 %)	660	675	650	665	2650
	B2 (1.0 %)	730	745	720	735	2930
	B3 (1.2 %)	740	755	730	750	2975

1. Cálculo de los totales marginales

Totales de A:

$A_1 = 2490 + 2730 + 2810 = 8030$

$A_2 = 2650 + 2930 + 2975 = \underline{8555}$

Totales de B:

$B_1 = 2490 + 2650 = \underline{5140}$

$B_2 = 2730 + 2930 = 5660$

$B_3 = 2810 + 2975 = \underline{5785}$

Total general:

$G = 8030 + 8555 = 16585$

2. Cálculo de sumas de cuadrados

$$C = \frac{16585^2}{24} = 11{,}460{,}926{,}04$$

$$\sum y_i^2 = 620^2 + 635^2 + \ldots + 750^2 = 11{,}506{,}825$$

Sumas de cuadrados:

$$SC_T = 11{,}506{,}825 - 11{,}460{,}926{,}04 = 45{,}898{,}96$$

$$SC_A = \frac{8030^2 + 8555^2}{3\times4} - C = \frac{64{,}481{,}600 + 73{,}188{,}025}{12} - 11{,}460{,}926{,}04 = 11{,}472{,}302{,}08 - 11{,}460{,}926{,}04 = 11{,}376{,}04$$

$$SC_B = \frac{5140^2 + 5660^2 + 5785^2}{2\times4} - C = \frac{26{,}419{,}600 + 32{,}035{,}600 + 33{,}466{,}225}{8} - = \frac{91{,}921{,}425}{8} - 11{,}460{,}926{,}04 = 11{,}490{,}178{,}13 - 11{,}460{,}926{,}04 = 29{,}252{,}09$$

$$SC_E = 45{,}898{,}96 - 11{,}376{,}04 - 29{,}252{,}09 = 5{,}043{,}75$$

3. Tabla ANOVA

Fuente	GL	SC	CM	F	p
Proteína (A)	1	11,376,04	11,376,04	40,58	< 0,001
Lisina (B)	2	29,252,09	14,626,05	52,17	< 0,001
A × B	2	227,08	113,54	0,41	> 0,05
Error	18	5,043,75	280,21		
Total	23	45,898,96			

4. Verificación de GL

$1 + 2 + 2 + 18 = 23$ ✓

5. Prueba de significancia (α = 0,05)

1. $F_{AB} = 0{,}41 < 3{,}55 \rightarrow$ No significativa (Como 0,41 < 3,55, la interacción no es significativa).
2. $F_A = 40{,}58 > 4{,}41 \rightarrow$ Proteína significativa. $F_B = 52{,}17 > 3{,}55 \rightarrow$ Lisina significativa. (La dieta con 20 % de proteína produce mayor GDP).
3. Aumentar la proteína y la lisina mejora la GDP independientemente (sin interacción).

Conclusión

Aumentar la proteína y la lisina mejora la **GDP** independientemente (sin interacción). El efecto de cada factor es consistente en todos los niveles del otro.

Figura 38. Caso Clínico.

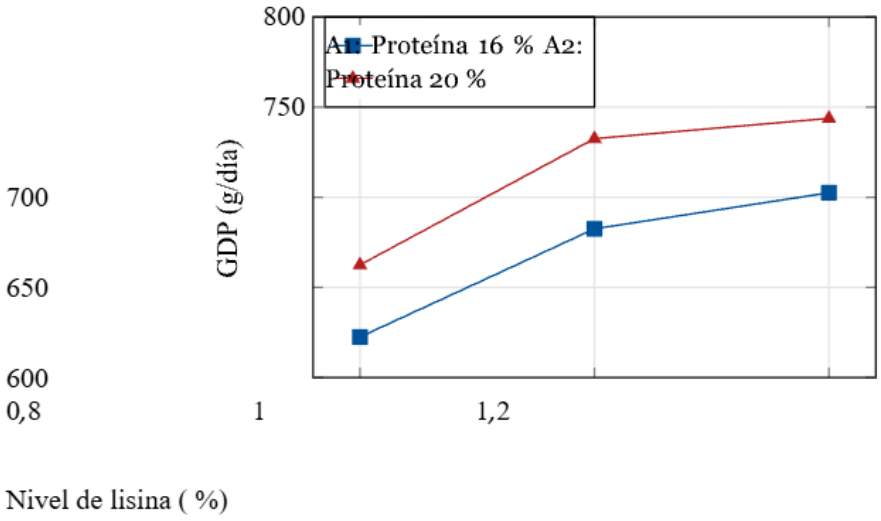


Figura 39. Gráfico de interacción para la factorial proteína lisina. Las líneas aproximadamente paralelas confirman la ausencia de interacción significativa: el efecto de la lisina es consistente en ambos niveles de proteína.

Aplicación Veterinaria

¿Qué habría pasado con dos experimentos separados? Si se hubiera evaluado solo la proteína (ignorando la lisina), el investigador habría necesitado un segundo experimento para evaluar la lisina. Con dos experimentos separados:
(1) se necesitan más animales en total, (2) no se puede detectar la interacción, y
(3) la variabilidad del error puede ser diferente entre experimentos, impidiendo comparaciones válidas. La factorial resuelve los tres problemas con un solo diseño (Box et al., 2005).

Ejemplo Resuelto

Ejemplo conceptual de interacción significativa.

Suponga que en un factorial 2 × 2 (antibiótico A vs. B × dosis baja vs. alta) se obtienen estas medias de concentración mínima inhibitoria (CMI):

	Dosis baja	Dosis alta
Antibiótico A	8.0	2.0
Antibiótico B	4.0	6.0

Medias marginales: $\bar{A} = (8 + 2)/2 = 5{,}0$, $\bar{B} = (4 + 6)/2 = 5{,}0$.

Efecto principal de antibiótico = 0 (¡las medias marginales son iguales!).

Pero: con dosis baja, B es mejor (CMI = 4 vs. 8); con dosis alta, A es mejor (CMI = 2 vs. 6). Los efectos principales son engañosos: hay interacción.

Efectos simples (correcto):

Efecto del antibiótico con dosis baja: B mejor que A (4 vs. 8).

Efecto del antibiótico con dosis alta: A mejor que B (2 vs. 6).

Recomendación: usar A con dosis alta, B con dosis baja.

Figura 40. Ejemplo Resuelto.

Ejemplo veterinario: evaluar el efecto del tipo de alojamiento (galpón abierto vs. cerrado) y el nivel de suplementación (3 niveles) sobre la producción de leche. El tipo de alojamiento solo puede asignarse a nivel de galpón completo (parcela principal), pero la suplementación puede asignarse a vacas individuales dentro del galpón (sub-parcela), disposición que representa la Figura 40. No es práctico tener la mitad de un galpón abierto y la otra mitad cerrado.

Concepto Clave

El diseño en **parcela dividida** (split-plot) se emplea cuando un factor es difícil o imposible de aleatorizar a nivel de la unidad experimental más pequeña, de tal manera que el factor "difícil" se aplica a unidades grandes (**parcela principal** o whole plot), mientras que el factor "fácil" se destina a subdivisiones dentro de esas unidades (**sub-parcela** o split plot) (Montgomery, 2017; Steel et al., 1997).

Alerta Importante

¿Cómo identificar cuándo se necesita parcela dividida? Pregúntese: "¿Pue-do aleatorizar *todos* los factores al mismo nivel?" Si un factor solo puede aplicarse a unidades grandes (galpones, fincas, estanques) y otro a unidades pequeñas dentro de ellas (animales, parcelas de pasto), necesita parcela dividida. Si ambos factores pueden aleatorizarse al mismo nivel, use una factorial simple (Lazic, 2016; Mead et al., 2012).

4.5.5 Estructura del diseño

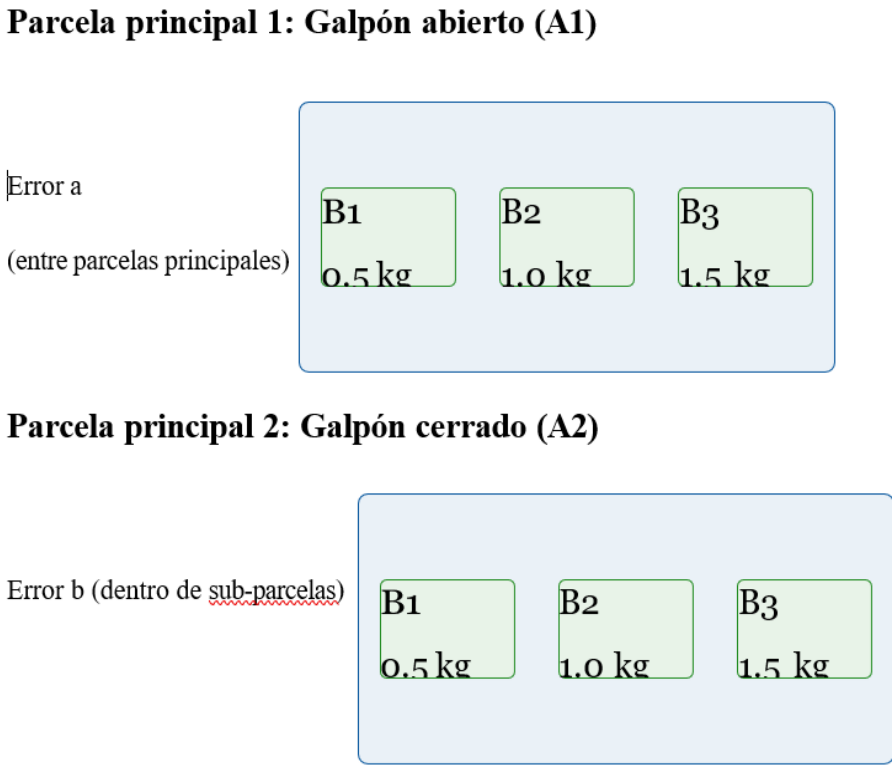


Figura 41. Estructura de un diseño en parcela dividida. El factor A (tipo de alojamiento) se aplica a parcelas principales; el factor B (suplementación) se aleatoriza dentro de cada parcela principal. Hay dos errores (a- entre parcelas) (b- dentro de sub-parcelas).

4.5.6 Modelo estadístico

Fórmula Clave

$$Y_{ijk} = \mu + \alpha_i + \delta_{ij} + \beta_k + (\alpha\beta)_{ik} + \varepsilon_{ijk} \quad (4.10)$$

Donde:

α_i: efecto del factor A (parcela principal)

δ_{ij}: error de parcela principal $(j = 1, \ldots, r)$, variación entre galpones con el mismo tipo de alojamiento

β_k: efecto del factor B (sub-parcela)

$(\alpha\beta)_{ik}$: interacción $A \times B$

ε_{ijk}: error de sub-parcela, variación entre vacas dentro del mismo galpón

Punto clave: El factor A se prueba contra δij (Error a), no contra εijk (Error b). El factor B y la interacción A x B se prueban contra εijk (Error b) (Steel et al., 1997; Winer et al., 1991).

Concepto Clave

¿Por qué dos errores? Porque la precisión con la que se estiman A y B es diferente: el factor A se aplica a unidades grandes (galpones), donde hay pocos galpones, mucha variación entre ellos y menos precisión. El factor B, en cambio, se aplica a unidades pequeñas (vacas dentro de galpones), en las que hay muchas vacas comparadas dentro del mismo galpón, con más precisión, de manera que usar un solo error sería incorrecto, como refleja la Tabla 23 al separar el Error a del Error b, ya que trataría ambos factores con la misma precisión cuando en realidad no la tienen (Montgomery, 2017).

4.5.7 Tabla ANOVA de la parcela dividida

Fuente	GL	SC	CM	*F*
Factor A (parcela ppal.)	$a-1$	SC_A	CM_A	CM_A/CM_δ
Error a (δ)	$a(r-1)$	SC_δ	CM_δ	
Factor B (sub-parcela)	$b-1$	SC_B	CM_B	CM_B/CM
$A \times B$	$(a-1)(b-1)$	SC_{AB}	CM_{AB}	
Error b (ε)	$a(r-1)(b-1)$	SC_ε	CM_ε	
Total	$abr-1$	SC_T		

Figura 42. Tabla ANOVA.

Alerta Importante

El error más frecuente en la parcela dividida es usar **el mismo error** para probar ambos factores, pese a que el factor A tiene menos precisión que el factor B porque se basa en menos unidades (parcelas principales). Consecuencias de usar el error equivocado:

Usar Error b para probar A → F demasiado grande → falsos positivos (inflación del error tipo I).

Usar Error a para probar B → F demasiado pequeño → falsos negativos (pérdida de potencia).

Siempre verifique que cada factor se prueba contra su error correspondiente (Lazic, 2016; Montgomery, 2017).

4.5.8 Ejemplo: alojamiento × suplementación en vacas lecheras

Caso Clínico

Se evalúan 2 tipos de alojamiento (A1: galpón abierto, A2: galpón cerrado) y 3 niveles de suplementación (B1: 0.5 kg, B2: 1.0 kg, B3: 1.5 kg de concentrado) sobre la producción de leche (L/día). Diseño: parcela dividida con r = 3 repeticiones (galpones) del factor A. Dentro de cada galpón, se asignan 3 vacas aleatoriamente a los 3 niveles de B.

Alojamiento	Rep.	B1 (0.5 kg)	B2 (1.0 kg)	B3 (1.5 kg)	Total parcela
A1 (Abierto)	1	22.0	25.5	27.0	74.5
A1 (Abierto)	2	20.5	24.0	26.5	71.0
A1 (Abierto)	3	21.5	25.0	26.0	72.5
A2 (Cerrado)	1	24.0	28.0	30.5	82.5
A2 (Cerrado)	2	23.5	27.5	29.0	80.0
A2 (Cerrado)	3	24.5	27.0	30.0	81.5

Figura 43. Caso Clínico.

Ejemplo Resuelto

1. Cálculo de totales, sumas de cuadrados y factores (primer bloque)

$a = 2,\ b = 3,\ r = 3,\ n = 18$

$G = 74{,}5 + 71{,}0 + 72{,}5 + 82{,}5 + 80{,}0 + 81{,}5 = 462{,}0$

$$\sum y^2 = \frac{462{,}0^2}{18} + \frac{213{,}444}{18} = 11{,}858{,}00$$

$$SCT = 12{,}046{,}50 - 11{,}858{,}00 = 188{,}50$$

Factor A (alojamiento):

$A_1 = 218{,}0,\ A_2 = 244{,}0$

$$SC_A = \frac{218{,}0^2 + 244{,}0^2}{9} - C = \frac{47{,}524 + 59{,}536}{9} - 11{,}858{,}00 = 11{,}895{,}56 - 11{,}858{,}00 = 37{,}56$$

Error a:

Totales de parcela principal: $P_{11} = 74{,}5,\ P_{12} = 71{,}0,\ P_{13} = 72{,}5,\ P_{21} = 82{,}5,\ P_{22} = 80{,}0,\ P_{23} = 81{,}5$

$$SC_{\text{parcelas}} = \frac{74{,}5^2 + 71{,}0^2 + 72{,}5^2 + 82{,}5^2 + 80{,}0^2 + 81{,}5}{3} - C = 11{,}858{,}00 - 11{,}898{,}67 = 40{,}67$$

$$SC_\delta = SC_{parcelas} - SC_T = 40{,}67 - 37{,}56 = 3{,}11$$

2. Factor B (suplementación)

$B_1 = 136{,}0,\ B_2 = 157{,}0,\ B_3 = 169{,}0$

$$SC_B = \frac{18{,}0 + 24{,}6{,}0 + 28{,}5{,}0}{3} - C = 11{,}989{,}00 - 11{,}858{,}00 = 93{,}00$$

3. Interacción A × B

Celdas $(AB)_{ij}$:

$AB_{11} = 64{,}0,\ AB_{12} = 74{,}5,\ AB_{13} = 79{,}5$

$AB_{21} = 72{,}0,\ AB_{22} = 82{,}5,\ AB_{23} = 89{,}5$

$$SC_{AB(\text{sub})} = \frac{64{,}0^2 + 74{,}5^2 + 79{,}5^2 + 72{,}0^2 + 82{,}5^2 + 89{,}5^2}{3} - C$$

$$= 11{,}989{,}00 - 11{,}858{,}00 = 131{,}00$$

$SC_{AB} = 0{,}44$

$SC_E = 54{,}39$

4. Tabla ANOVA

Fuente	GL	SC	CM	F	p
Alojamiento (A)	1	37,56	37,56	48,28	< 0,01
Error a	4	3,11	0,778	-	-
Suplementación (B)	2	93,00	46,50	6,84	< 0,05
A × B	2	0,44	0,22	0,03	> 0,05
Error b	8	54,39	6,80	-	-
Total	17	188,50	-	-	-

5. Verificación de GL

$1 + 4 + 2 + 2 + 8 = 17$ ✓

6. Decisiones

1. A: $F = 37{,}56/0{,}778 = 48{,}28$, $F_{tab}(1, 4) = 7{,}71$ → significativo ($p < 0{,}01$).
   El galpón cerrado produce más leche.
2. B: $F = 46{,}50/6{,}80 = 6{,}84$, $F_{tab}(2, 8) = 4{,}46$ → significativo ($p < 0{,}05$).
   Mayor suplementación aumenta la producción.
3. $A \times B$: $F = 0{,}22/6{,}80 = 0{,}03$ → no significativo.
   Los efectos son aditivos.

Figura 44. Ejercio Resuelto.

Nota: Observe en la Tabla 24 que CM_{δ} = 0,778 (Error a) es mucho menor que CM_{E} = 6,80 (Error b). Esto indica que la variación entre galpones del mismo tipo es pequeña, pero la variación entre vacas dentro de galpones es considerable. El factor A se prueba con alta precisión (Error a pequeño); el factor B con menor precisión (Error b grande) (Dohoo et al., 2009).

4.6 Diseños especiales: visión panorámica

Los siguientes diseños se presentan conceptualmente para que el estudiante los reconozca y sepa cuándo consultar textos avanzados. Su análisis cuantitativo excede el alcance de este libro.

4.6.1 Medidas repetidas

Concepto Clave

En un diseño de **medidas repetidas**, la misma UE se mide en múltiples ocasiones (tiempos, condiciones), lo cual aporta mayor potencia al usar cada sujeto como su propio control, si bien el desafío es que las observaciones dentro del mismo sujeto no son independientes, lo que viola un supuesto del ANOVA estándar y requiere correcciones (Greenhouse-Geisser, Huynh-Feldt) o modelos mixtos (Girden, 1992; Pinheiro & Bates, 2000).

Ejemplo: medir la glucemia de un mismo grupo de perros a las 0, 2, 4, 6 y 8 horas post-prandial, donde la variable tiempo tiene 5 niveles, aunque las 5 mediciones no son independientes, ya que, si

un perro tiene glucemia basal alta, probablemente la tendrá alta en todos los tiempos.

4.6.2 Diseño cruzado (crossover)

En un diseño **crossover**, cada sujeto recibe todos los tratamientos en secuencia, con un período de lavado entre ellos. Es un caso especial de medidas repetidas muy usado en farmacología veterinaria (Dohoo et al., 2009; Thrusfield, 2018).

Ejemplo: crossover 2 x 2 para comparar dos formulaciones de ivermectina en bovinos: el grupo 1 recibe la formulación A en el período 1 y la B en el período 2; el grupo 2 recibe B primero y A después.

Alerta Importante

En los diseños crossover se debe evaluar el **efecto de arrastre** (carryover): el efecto residual de un tratamiento que persiste en el período siguiente. Si el período de lavado es insuficiente, el efecto de arrastre confunde los resultados. Esto es especialmente relevante en fármacos de larga vida media (Petrie & Watson, 2013).

4.6.3 Diseños anidados vs. cruzados

Tabla 16. Comparación entre diseños anidados y cruzados.

Característica	Anidado	Cruzado (factorial)
Niveles de B	Diferentes dentro de cada A	Los mismos en todos los A
Ejemplo	Veterinarios dentro de clínicas; dietas dentro de razas	—
Interacción	No se puede estimar	Se puede estimar
Análisis	ANOVA jerárquica	ANOVA factorial

4.6.4 Diseños aumentados

Los diseños **aumentados** permiten evaluar un gran número de tratamientos nuevos (sin replicar) junto con testigos replicados. Son útiles en evaluación preliminar de genotipos o compuestos. El error se estima solo de los testigos (Casella, 2008; Oehlert, 2010).

4.7 Guía de selección de diseño

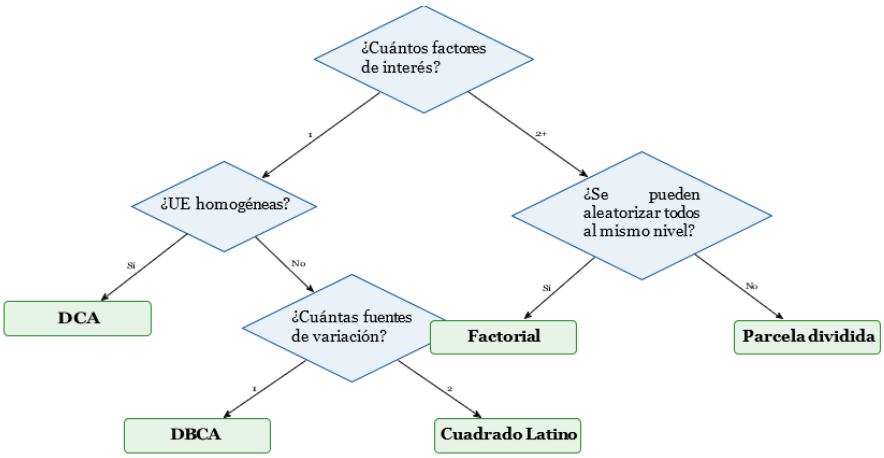


Figura 45. Flujograma de selección de diseño experimental. La primera decisión es si se evalúa uno o más factores; la segunda, la estructura de variabilidad o las

restricciones de aleatorización.

Tabla 17. Resumen comparativo de todos los diseños experimentales del libro.

Diseño	Factores	Bloques	Interacción	Uso típico
DCA	1	0	—	Laboratorio, UE homogéneas
DBCA	1	1	—	Campo, producción
Cuadrado Latino	1	2	—	—
Crossover, Factorial	≥ 2	0–1	Sí	Nutrición, farmacología
Parcela dividida	≥ 2	0–1	Sí	Alojamiento × dieta

Aplicación Veterinaria

¿Cómo elegir el diseño en la práctica?

Hágase estas preguntas en orden, recogidas también en el flujograma de la aplicación.

1 ¿Cuántos factores quiero evaluar?

Si 1 → DCA/DBCA/DCL. Si ≥ 2 → Factorial/Parcela dividida.

Diseño	Factores	Bloques	Interacción	Uso típico
DCA	1	0	—	Laboratorio, UE homogéneas
DBCA	1	1	—	Campo, producción
Cuadrado Latino	1	2	—	—
Crossover, Factorial	≥ 2	0–1	Sí	Nutrición, farmacología
Parcela dividida	≥ 2	0–1	Sí	Alojamiento × dieta

2 ¿Mis UE son homogéneas?

Si sí → DCA. Si no → DBCA (1 fuente) o DCL (2 fuentes).

3 ¿Puedo aleatorizar todos los factores al mismo nivel?

Si sí → Factorial. Si no → Parcela dividida.

4 ¿Cuántas réplicas necesito?

Use la ecuación de recursos ($10 \leq E \leq 20$).

5 ¿Es ético usar este número de animales?

Consulte las guías ARRIVE y las 3Rs
(Percie Perçie du Sert et al., 2020; Kilkenny et al., 2010; National Research Council, 2011).

Figura 46. Aplicación Veterinaria.

4.8 Conceptos clave

- Un **experimento factorial** evalúa todas las combinaciones de niveles de dos o más factores simultáneamente.

- La **interacción** ocurre cuando el efecto de un factor depende del nivel del otro; se detecta con líneas no paralelas en el gráfico de interacción.

- Cuando hay interacción significativa, los efectos principales no se interpretan directamente; se analizan los **efectos simples**.

- La **parcela dividida** tiene dos errores: Error a para el factor de parcela principal (menos preciso) y Error b para el factor de sub-parcela (más preciso).

- La selección del diseño depende del número de factores, la estructura de variabili-dad y las restricciones de aleatorización, criterios que la Tabla 26 compara para todos los diseños del libro.

- Siempre evalúe la interacción *antes* de interpretar los efectos principales.

- El factorial es más eficiente que dos experimentos separados: usa menos UE y detecta interacciones.

- La SC_{AB} se calcula indirectamente: es lo que "sobra" después de extraer SC_A y SC_B de la SC de subtotales.

4.9 Actividades de consolidación

1. **Actividad 1:** Diseñe un factorial 3 × 2 para evaluar 3 antibióticos (amoxicilina, enrofloxacina, florfenicol) y 2 vías de administración (oral, inyectable) sobre la eficacia clínica en aves con colibacilosis. Defina: factores, niveles, combinaciones, UE, réplicas. Escriba la estructura de la tabla ANOVA con los GL.

2. **Actividad 2:** Un investigador observa que la suplementación con zinc mejora la ganancia de peso en pollos con dieta baja en proteína, aunque no tiene efecto con dieta alta en proteína; ¿hay interacción? Dibuje el gráfico de interacción esperado y explique por qué los efectos principales serían engañosos.

3. **Actividad 3:** Proponga un diseño en parcela dividida para un ensayo que evalúe 2 sistemas de pastoreo (rotacional vs. continuo) y 3 razas de ovinos sobre la producción de lana, identifique con la debida justificación cuál factor es la parcela principal y cuál la sub-parcela, y escriba la tabla ANOVA con GL.

4. **Actividad 4:** A partir de la tabla ANOVA del ejemplo de nutrición porcina (factorial 2 × 3), aplique la prueba de Tukey al factor B (lisina) para determinar cuáles niveles difieren. Use $q_{0,05,3,18} = 3{,}609$. Medias: $\bar{B}_1 = 642{,}5$, $\bar{B}_2 = 707{,}5$, $\bar{B}_3 = 723{,}1$.

4.10 Ejercicios propuestos

1. En un factorial 2 x 2 con 5 réplicas, ¿cuántas UE se necesitan? ¿Cuántos GL tiene el error? Respuesta: 20 UE, glE = 16.

2. Un factorial 3 x 4 en DBCA con 3 bloques tiene SCT = 500, SCA = 120, SCB = 200, SCAB = 45, SCBloq = 80. Calcule SCE y los GL de cada fuente. Respuesta: SCE = 55, glE = 16.

3. ¿Cuál es la diferencia entre un efecto principal y un efecto simple? ¿Cuándo se prefiere uno sobre otro?

4. En un diseño en parcela dividida con a = 2, b = 4, r = 3, ¿cuántos GL tiene el Error a? ¿Y el Error b? Respuesta: Error a: 4, Error b: 12.

5. Un investigador analiza un factorial 2x 3 y obtiene FA×B = 5,2 con p = 0,02.

6. ¿Tiene sentido interpretar los efectos principales de A y B individualmente? Justifique.

7. Diseñe un experimento que evalúe el efecto de la temperatura del agua (20, 25, $^{3}0$ $^{\circ}$C) y de la densidad de siembra (10, 20, 30 peces/m3) sobre el crecimiento de tilapia, e indique qué diseño resulta apropiado: ¿puede resolverse con una factorial simple o se necesita parcela dividida?

8. En una parcela dividida, ¿por qué el factor de parcela principal

se prueba con Error a y no con Error b? Explique en términos de la estructura de aleatorización.

9. Un ensayo factorial 2x 2 evalúa sexo (macho, hembra) y dieta (control, expe-rimental) en conejos. ¿Es el sexo un factor "asignado" o "manipulado"? ¿Qué implicación tiene para la interpretación causal?

10. Construya una tabla ANOVA vacía para un factorial 3 x 2 en DCA con 5 réplicas. Indique los GL de cada fuente. Respuesta: A: 2, B: 1, AB: 2, Error: 24, Total: 29.

11. Explique conceptualmente qué es un diseño crossover 2x 2 y por qué se necesita un período de lavado. Dé un ejemplo veterinario.

12. Se evaluó un factorial 2x 2 (raza: Holstein vs. Jersey dieta: estándar vs. alta energía) con r = 6 vacas por combinación.

Tabla 18. Los datos de producción de leche (L/día).

Raza	**Dieta**	**Medias de celda**
Holstein	Estándar	28.5
Holstein	Alta energía	34.2
Jersey	Estándar	22.0
Jersey	Alta energía	24.5

¿Hay interacción? Justifique observando las diferencias. *Respuesta: Sí. La dieta alta mejora 5.7 L en Holstein pero solo 2.5 L en Jersey.*

13. Un investigador quiere evaluar 3 niveles de un fármaco y 2

vías de administración en equinos. Si usa una factorial 3 2 con r = 4, ¿cuántos caballos necesita? ¿Es ético? Use la ecuación de recursos. Respuesta: 24 caballos, E = 18, adecuado.

4.11 Preguntas de autoevaluación

Selección múltiple (la respuesta correcta se indica con ←):

1. En una factorial 2 × 3, el número de combinaciones de tratamientos es:

a) 2.

b) 3.

c) 5.

d) 6. ←

2. La interacción $A \times B$ se detecta gráficamente cuando:

a) Las líneas son paralelas.

b) Las líneas se cruzan o divergen. ←

c) Ambos factores son significativos.

d) El error es muy pequeño.

3. En presencia de interacción significativa, se deben interpretar:

a) Solo los efectos principales.

b) Solo el error.

c) Los efectos simples. ←

d) Los bloques.

4. En una parcela dividida, el factor de parcela principal se prueba contra:

a) El Error b (sub-parcela).

b) El Error a (parcela principal). ←

c) El CM total.

d) El CM de tratamientos.

5. La ventaja principal de un factorial sobre dos experimentos separados es:

a) Usa menos estadísticos.

b) Detecta interacciones y es más eficiente. ←

c) No necesita aleatorización.

d) Siempre tiene menos error.

6. En un factorial 3 × 2 con 4 réplicas en DCA, los GL del error son:

a) 18. ←

b) 23.

c) 12.

d) 20.

7. Un diseño en parcela dividida se necesita cuando:

a) Todos los factores se aleatorizan al mismo nivel.

b) Un factor solo puede aplicarse a unidades grandes. ←

c) No hay interacción.

d) Se trabaja con un solo factor.

8. En un diseño crossover, el período de lavado sirve para:

a) Aumentar el tamaño de muestra.

b) Eliminar el efecto residual del tratamiento anterior. ←

c) Aleatorizar los tratamientos.

d) Bloquear por animal.

9. Si $SC_{AB} = 0$, esto indica:

a) Los factores no tienen efecto.

b) No hay interacción (efectos aditivos). ←

c) Los efectos principales son iguales.

d) Se cometió un error de cálculo.

10. La tabla ANOVA de una parcela dividida tiene:

a) Un solo error.

b) Dos errores (Error a y Error b). ←

c) Tres errores.

d) Ningún error.

11. En un factorial, la SC_{AB} se calcula como:

a) $SC_A \times SC_B$.

b) $SC_{\text{subtotales}} - SC_A - SC_B$. ←

c) $SC_T - SC_E$.

d) $SC_A + SC_B$.

12. Cuando NO hay interacción significativa:

a) No se puede interpretar nada.

b) Se interpretan los efectos principales directamente. ←

c) Se ignoran los efectos principales.

d) Se repite el experimento.

Preguntas de desarrollo:

1. Explique con un ejemplo veterinario concreto la diferencia entre un efecto principal y una interacción. ¿Por qué la interacción es biológicamente importante en la investigación con

animales?

2. Compare la factorial completa y el diseño en parcela dividida: ¿cuándo se prefiere cada uno? ¿Cuál tiene mayor potencia para detectar efectos del factor B (sub-parcela)? Justifique.

3. A partir de todo lo estudiado en este libro (DCA, DBCA, Cuadrado Latino, Factorial, Parcela Dividida), diseñe un estudio completo que evalúe el efecto de 2 razas de pollos y 3 niveles de proteína sobre la conversión alimenticia, definiendo diseño, modelo, factores, niveles, UE, réplicas y variable respuesta, para escribir después la estructura completa de la tabla ANOVA con GL.

4.12 Glosario

Combinación de tratamientos

Cada par específico de niveles de los factores en una factorial (ej.: A_1B_2).

Crossover (cruzado)

Diseño donde cada sujeto recibe todos los tratamientos en secuencia con períodos de lavado.

Diseño anidado Diseño donde los niveles de un factor están contenidos dentro de los niveles de otro (no cruzados).

Diseño aumentado

Diseño que evalúa muchos tratamientos sin réplica junto con testigos replicados.

Efecto de arrastre (carryover)

Efecto residual de un tratamiento que persiste en el período siguiente en un diseño crossover.

Efecto principal

Efecto promedio de un factor sobre todos los niveles del otro factor.

Efecto simple Efecto de un factor en un nivel fijo del otro factor; se usa cuando hay interacción significativa.

Efectos aditivos

Cuando no hay interacción: el efecto combinado de A y B es la suma de sus efectos individuales.

Error a (parcela principal)

Error experimental para probar el factor de parcela principal en un split-plot.

Error b (sub-parcela)

Error experimental para probar el factor de sub-parcela y la interacción en un split-plot.

Experimento factorial

Diseño que evalúa todas las combinaciones de niveles de dos o más

factores.

Factorial completo

Factorial que incluye todas las combinaciones posibles (a diferencia del fraccionado).

Gráfico de interacción

Gráfico de medias de celda que muestra el efecto de un factor en cada nivel del otro; líneas paralelas = sin interacción.

Interacción (A x B)

Efecto que ocurre cuando la respuesta a un factor depende del nivel del otro.

Medias marginales

Medias de un factor promediadas sobre todos los niveles del otro; base de los efectos principales.

Medidas repetidas

Diseño donde la misma UE se mide en múltiples ocasiones (tiempos, condiciones).

Parcela dividida (split-plot)

Diseño factorial con restricción en la aleatorización: un factor se aplica a unidades grandes y otro a subdivisiones.

Parcela principal (whole plot)

Unidad grande a la que se aplica el factor difícil de aleatorizar.

Período de lavado (washout)

Tiempo entre tratamientos en un crossover para eliminar efectos residuales.

Sub-parcela (split plot)

Subdivisión dentro de la parcela principal donde se aplica el factor fácil de aleatorizar.

Subtotales de celda

Totales de las observaciones en cada combinación $(AB)_{ij}$; necesarios para calcular SC_{AB}.

4.13 Resumen

Los **experimentos factoriales** permiten evaluar simultáneamente dos o más factores y detectar interacciones entre ellos, algo imposible con diseños un factoriales, y la interacción, que se verifica cuando el efecto de un factor depende del nivel del otro, constituye un fenómeno biológicamente común en investigación veterinaria y de notable relevancia práctica; cuando no es significativa, los efectos principales se interpretan directamente, mientras que, si lo es, se deben analizar los efectos simples, ya que las medias marginales pueden ser engañosas (Box et al., 2005; Montgomery, 2017).

El **diseño en parcela dividida** resuelve el problema de factores con diferente facilidad de aleatorización mediante la introducción de

dos niveles de error: el Error a para el factor de parcela principal (menos preciso, basado en pocas unidades grandes) y el Error b para el factor de sub-parcela (más preciso, basado en muchas unidades pequeñas), no obstante lo cual el error más frecuente consiste en usar un solo error para ambos factores, lo que produce conclusiones incorrectas (Steel et al., 1997; Winer et al., 1991).

Otros diseños (medidas repetidas, crossover, anidados, aumentados) amplían el repertorio del investigador veterinario para situaciones especiales (Casella, 2008; Girden, 1992; Pinheiro & Bates, 2000), y la selección del diseño apropiado depende a su vez del número de factores, la estructura de variabilidad, las restricciones de aleatorización y los principios éticos de uso de animales (Kilkenny et al., 2010; Russell & Burch, 1959).

Este libro ha presentado las herramientas fundamentales del diseño experimental: desde los principios de Fisher (réplica, aleatorización, bloqueo) hasta los diseños clásicos (DCA, DBCA, Cuadrado Latino) y los factoriales, pasando por la prueba de hipótesis y el ANOVA, de tal manera que, con estas bases, el estudiante de medicina veterinaria está preparado para planificar investigaciones rigurosas, analizar datos con propiedad e interpretar críticamente la literatura científica de su disciplina.

REFERENCIAS BIBLIOGRÁFICAS

Altman, D. G. (1980). Statistics and Ethics in Medical Research: III. How Large a Sample? British Medical Journal, 281(6251), 1336-1338. https://doi.org/10.1136/bmj.281.6251.1336

Arifin, W. N., & Zahiruddin, W. M. (2017). Sample Size Calculation in Animal Studies Using Resource Equation Approach. Malaysian Journal of Medical Sciences, 24(5), 101-105. https://doi.org/10.21315/mjms2017.24.5.11

Bartlett, M. S. (1937). Properties of Sufficiency and Statistical Tests. Proceedings of the Royal Society of London. Series A, 160(901), 268-282. https://doi.org/10.1098/rspa.1937.0109

Bate, S. T., & Clark, R. A. (2014). The Design and Statistical Analysis of Animal Experiments. Cambridge University Press.

Bonferroni, C. E. (1936). Teoria statistica delle classi e calcolo delle probabilità. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze, 8, 3-62.

Box, G. E. P., Hunter, J. S., & Hunter, W. G. (2005). Statistics for Experimenters: Design, Innovation, and Discovery (2.a ed.). John Wiley & Sons.

Casella, G. (2008). Statistical Design. Springer.

Charan, J., & Kantharia, N. D. (2013). How to Calculate Sample Size in Animal Studies? Journal of Pharmacology and

Pharmacotherapeutics, 4(4), 303-306. https://doi.org/10.4103/0976-500X.119726

Cochran, W. G., & Cox, G. M. (1957). Experimental Designs (2.a ed.). John Wiley & Sons.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2.a ed.). Lawrence Erlbaum Associates.

Cumming, G. (2009). Inference by Eye: Reading the Overlap of Independent Confidence Intervals. Statistics in Medicine, 28(2), 205-220. https://doi.org/10.1002/sim.3471

Daniel, W. W., & Cross, C. L. (2013). Biostatistics: A Foundation for Analysis in the Health Sciences (10.a ed.). John Wiley & Sons.

Dell, R. B., Holleran, S., & Ramakrishnan, R. (2002). Sample Size Determination. ILAR Journal, 43(4), 207-213. https://doi.org/10.1093/ilar.43.4.207

Dohoo, I. R., Martin, S. W., & Stryhn, H. (2009). Veterinary Epidemiologic Research (2.a ed.). VER Inc.

Duncan, D. B. (1955). Multiple Range and Multiple F Tests. Biometrics, 11(1), 1-42. https://doi.org/10.2307/3001478

Dunnett, C. W. (1955). A Multiple Comparison Procedure for Comparing Several Treatments with a Control. Journal of the American Statistical Association, 50(272), 1096-1121. https://doi.org/10.1080/01621459.1955.10501294

Festing, M. F. W. (2006). Design and Statistical Methods in Studies Using Animal Models of Development. ILAR Journal, 47(1), 5-14. https://doi.org/10.1093/ilar.47.1.5

Festing, M. F. W. (2018). On Determining Sample Size in Experiments Involving Laboratory Animals. Laboratory Animals, 52(4), 341-350. https://doi.org/10.1177/0023677217738268

Festing, M. F. W., & Altman, D. G. (2002). Guidelines for the Design and Statistical Analysis of Experiments Using Laboratory Animals. ILAR Journal, 43(4), 244-258. https://doi.org/10.1093/ilar.43.4.244

Festing, M. F. W., Overend, P., Das, R. G., Borja, M. C., & Berdoy, M. (2002). The Design of Animal Experiments: Reducing the Use of Animals in Research Through Better Experimental Design. Royal Society of Medicine Press.

Fisher, R. A. (1925). Statistical Methods for Research Workers. Oliver & Boyd.

Fisher, R. A. (1935). The Design of Experiments. Oliver & Boyd.

Girden, E. R. (1992). ANOVA: Repeated Measures. SAGE Publications.

Hooijmans, C. R., Rovers, M. M., de Vries, R. B. M., Leenaars, M., Ritskes-Hoitinga, M., & Langendam, M. W. (2014). SYRCLE's Risk of Bias Tool for Animal Studies. BMC Medical Research Methodology, 14, 43. https://doi.org/10.1186/1471-2288-14-43

Hurlbert, S. H. (1984). Pseudoreplication and the Design of Ecological Field Experiments. Ecological Monographs, 54(2), 187-211. https://doi.org/10.2307/1942661

Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M., & Altman, D. G. (2010). Improving Bioscience Research Reporting: The ARRIVE Guidelines for Reporting Animal Research. PLoS Biology, 8(6), e1000412. https://doi.org/10.1371/journal.pbio.1000412

Kruskal, W. H., & Wallis, W. A. (1952). Use of Ranks in One-Criterion Variance Analysis. Journal of the American Statistical Association, 47(260), 583-621. https://doi.org/10.1080/01621459.1952.10483441

Kuehl, R. O. (2000). Design of Experiments: Statistical Principles of Research Design and Analysis (2.a ed.). Duxbury Press.

Lazic, S. E. (2010). The Problem of Pseudoreplication in Neuroscientific Studies. BMC Neuroscience, 11, 5. https://doi.org/10.1186/1471-2202-11-5

Lazic, S. E. (2016). Experimental Design for Laboratory Biologists: Maximising Information and Improving Reproducibility. Cambridge University Press.

Levene, H. (1960). Robust Tests for Equality of Variances. En I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow & H. B. Mann (Eds.), Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling (pp. 278-292). Stanford University Press.

Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. The Annals of Mathematical Statistics, 18(1), 50-60. https://doi.org/10.1214/aoms/1177730491

Mead, R., Gilmour, S. G., & Mead, A. (2012). Statistical Principles for the Design of Experiments. Cambridge University Press.

Montgomery, D. C. (2017). Design and Analysis of Experiments (9.a ed.). John Wiley & Sons.

Nakagawa, S., & Cuthill, I. C. (2007). Effect Size, Confidence Interval and Statistical Significance: A Practical Guide for Biologists. Biological Reviews, 82(4), 591-605. https://doi.org/10.1111/j.1469-185X.2007.00027.x

National Research Council. (2011). Guide for the Care and Use of Laboratory Animals (8.a ed.). https://doi.org/10.17226/12910

Oehlert, G. W. (2010). A First Course in Design and Analysis of Experiments. W. H. Freeman.

Ott, R. L., & Longnecker, M. T. (2016). An Introduction to Statistical Methods and Data Analysis (7.a ed.). Cengage Learning.

Pagano, M., & Gauvreau, K. (2018). Principles of Biostatistics (2.a ed.). CRC Press.

Percie du Sert, N., Hurst, V., Ahluwalia, A., Alam, S., Avey, M. T., Baker, M., Browne, W. J., Clark, A., Cuthill, I. C., Dirnagl, U., Emerson, M., Garner, P., Holgate, S. T., Howells, D. W., Karp, N.

A., Lazic, S. E., Lidster, K., MacCallum, C. J., Macleod, M., ... Würbel, H. (2020). The ARRIVE Guidelines 2.0: Updated Guidelines for Reporting Animal Research. PLoS Biology, 18(7), e3000410. https://doi.org/10.1371/journal.pbio.3000410

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences. Behavior Research Methods, 39(2), 175-191. https://doi.org/10.3758/BF03193146

Petrie, A., & Watson, P. (2013). Statistics for Veterinary and Animal Science (3.a ed.). Wiley-Blackwell.

Pinheiro, J. C., & Bates, D. M. (2000). Mixed-Effects Models in S and S-PLUS. Springer.

Quinn, G. P., & Keough, M. J. (2002). Experimental Design and Data Analysis for Biologists. Cambridge University Press.

Russell, W. M. S., & Burch, R. L. (1959). The Principles of Humane Experimental Technique. Methuen.

Scheffé, H. (1953). A Method for Judging All Contrasts in the Analysis of Variance. Biometrika, 40(1–2), 87-110. https://doi.org/10.1093/biomet/40.1-2.87

Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). Biometrika, 52(3–4), 591-611. https://doi.org/10.1093/biomet/52.3-4.591

Snedecor, G. W., & Cochran, W. G. (1989). Statistical Methods (8.a

ed.). Iowa State University Press.

Steel, R. G. D., Torrie, J. H., & Dickey, D. A. (1997). Principles and Procedures of Statistics: A Biometrical Approach (3.a ed.). McGraw-Hill.

Student. (1908). The Probable Error of a Mean. Biometrika, 6(1), 1-25. https://doi.org/10.2307/2331554

Thrusfield, M. (2018). Veterinary Epidemiology (4.a ed.). Wiley-Blackwell.

Tukey, J. W. (1949). Comparing Individual Means in the Analysis of Variance. Biometrics, 5(2), 99-114. https://doi.org/10.2307/3001913

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA Statement on p-Values: Context, Process, and Purpose. The American Statistician, 70(2), 129-133. https://doi.org/10.1080/00031305.2016.1154108

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). Statistical Principles in Experimental Design (3.a ed.). McGraw-Hill.

Yates, F. (1934). The Analysis of Multiple Classifications with Unequal Numbers in the Different Classes. Journal of the American Statistical Association, 29(185), 51-66. https://doi.org/10.1080/01621459.1934.10502686

Zar, J. H. (2010). Biostatistical Analysis (5.a ed.). Pearson Prentice Hall.

Diseño experimental en medicina veterinaria Principios, diseños clásicos y aplicaciones en investigación de especies, se publicó en el mes de noviembre de 2026.

ISBN: 978-9907-852-10-3

Grupo Editorial BLR
Ecuador
Cel: +593 98 320 4362
https://grupoblr.com/
publicaciones@grupoblr.com

BIOGRAFÍA DE AUTORES

Luiggi Oscar Solano Maza:
Es Bioquímico Farmacéutico con doble maestría en laboratorio clínico y diagnóstico. Docente e investigador en la Universidad Técnica de Machala, colabora en la Transformación Digital de la Academia, con líneas de investigación en diagnóstico clínico e inteligencia artificial aplicada a la educación y al laboratorio.

Ronald Andreé Vitonera Rogel:
Médico veterinario zootecnista, magíster en Análisis Biológico y Diagnóstico de Laboratorio y en Clínica y Cirugía de Pequeñas Especies. Docente de Fisiología Animal en la Universidad Técnica de Machala y candidato doctoral en la Universidad Autónoma de Madrid. Investiga resistencia antimicrobiana bajo el enfoque Una Sola Salud.

DISEÑO EXPERIMENTAL EN MEDICINA VETERINARIA PRINCIPIOS, DISEÑOS CLÁSICOS Y APLICACIONES EN INVESTIGACIÓN DE ESPECIES

Estimado lector, en el contexto actual de los sistemas de salud, caracterizados por el incremento de enfermedades crónicas y patologías complejas, la enfermería especializada desempeña un papel esencial en la prestación de cuidados integrales, seguros y humanizados. La creciente demanda asistencial exige profesionales con competencias científicas, clínicas y éticas que permitan abordar eficazmente las necesidades del paciente mediante el razonamiento clínico, la toma de decisiones basada en evidencia y la aplicación sistemática del Proceso de Atención de Enfermería como herramienta fundamental para garantizar una atención individualizada y de calidad.

La obra Enfermería Especializada: Fundamentos, Procesos y Cuidados Integrales en Patologías Complejas surge como una respuesta académica a estas necesidades formativas, ofreciendo un abordaje estructurado de patologías prioritarias que impactan significativamente la morbimortalidad hospitalaria. El texto integra fundamentos teóricos actualizados, fisiopatología, medios diagnósticos, tratamiento y cuidados enfermeros sustentados en las taxonomías NANDA, NIC y NOC, favoreciendo la articulación entre el conocimiento científico y la práctica clínica.

Dirigido a estudiantes y profesionales de enfermería, este libro constituye una herramienta pedagógica orientada al fortalecimiento de competencias clínicas en el manejo de pacientes con condiciones complejas, promoviendo la detección temprana de complicaciones, la monitorización continua, la educación terapéutica y el cuidado centrado en la persona. Su enfoque didáctico, apoyado en organizadores gráficos y casos clínicos, facilita el aprendizaje significativo y el desarrollo del juicio clínico enfermero.

En un escenario sanitario dinámico y en constante evolución, esta obra representa un aporte relevante para la formación de profesionales comprometidos con la práctica basada en evidencia, la seguridad del paciente y la excelencia del cuidado, consolidando el rol de la enfermería especializada como eje fundamental en la atención integral de la salud.

GRUPO BLR

Grupo Editorial BLR
Ecuador
Cel: +593 98 320 4362
https://grupoblr.com/
publicaciones@grupoblr.com

ISBN: 978-9907-852-10-3

UNIVERSIDAD TÉCNICA DE MACHALA
1969