



UEB
UNIVERSIDAD
ESTATAL DE BOLIVAR

BIOESTADÍSTICA APLICADA A LAS CIENCIAS VETERINARIAS

Toma decisiones basadas en evidencia animal cuantificable.

Víctor Alejandro Bosquez Barcenes
Joscelito Bolívar Solano Gaibor

ISBN: 978-9907-0-0420-5

2025

BIOESTADÍSTICA APLICADA A LAS CIENCIAS VETERINARIAS

AUTORES:

**VÍCTOR ALEJANDRO BOSQUEZ BARCENES
JOSCELITO BOLÍVAR SOLANO GAIBOR**

ISBN: 978-9907-0-0420-5



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 Estatal de Bolívar
Riobamba – Ecuador
Correo: publicaciones@grupobl.com
<https://grupobl.com/libros-investig>
REPOSITORIO



Bosquez, V., Solano, J. (2025) Bioestadística aplicada a las ciencias veterinarias. Grupo Editorial BLR.

© Víctor Alejandro Bosquez Barcenos
Joscelito Bolívar Solano Gaibor

ISBN: 978-9907-0-0420-5

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.

FILIACIONES DE LOS AUTORES

Víctor Alejandro Bosquez Barcenas

Universidad Estatal de Bolívar

Correo Electrónico: abosquez@ueb.edu.ec

ORCID: <https://orcid.org/0000-0001-7679-6023>

Joscelito Bolívar Solano Gaibor

Universidad Estatal de Bolívar

Correo Electrónico: jsolano@ueb.edu.ec

ORCID: <https://orcid.org/0000-0002-5109-2713>



PRÓLOGO

El presente libro surge de la necesidad imperante de proporcionar a los estudiantes de Medicina Veterinaria y Zootecnia una obra integral que aborde la estadística y el diseño experimental desde una perspectiva genuinamente aplicada a las ciencias veterinarias. Durante décadas, la enseñanza de la bioestadística en las facultades veterinarias ha dependido de textos genéricos que, si bien sólidos en sus fundamentos teóricos, carecen de la contextualización específica que requiere nuestra disciplina. Esta obra busca llenar ese vacío, ofreciendo un tratamiento riguroso de los métodos estadísticos mientras mantiene una conexión constante con la realidad de la práctica veterinaria y la producción animal.

La estadística en las ciencias veterinarias trasciende el mero ejercicio académico. En la práctica clínica diaria, el médico veterinario debe interpretar pruebas diagnósticas considerando sensibilidad y especificidad, establecer intervalos de referencia para diferentes especies y razas, y evaluar la eficacia de tratamientos basándose en evidencia cuantificable. En el ámbito de la producción animal, la optimización de parámetros productivos, el mejoramiento genético y la evaluación económica de sistemas de producción dependen fundamentalmente del análisis estadístico riguroso. En salud pública veterinaria, la vigilancia epidemiológica, el control de zoonosis y la evaluación de riesgos requieren un dominio profundo de métodos estadísticos aplicados.

La estructura de este libro refleja una progresión pedagógica cuidadosamente diseñada. Los primeros capítulos establecen los fundamentos de la bioestadística y la estadística descriptiva, proporcionando las bases necesarias para comprender conceptos más complejos. Los capítulos intermedios abordan la inferencia estadística, las pruebas paramétricas y no paramétricas, y el análisis de correlación y regresión, siempre con ejemplos extraídos de situaciones reales en medicina veterinaria. Los capítulos finales se adentran en el diseño experimental y técnicas avanzadas, preparando al estudiante para la investigación científica rigurosa.

Una característica distintiva de esta obra es su enfoque en la implementación práctica mediante software estadístico, particularmente R, que se ha convertido en el estándar de facto en investigación biomédica. Cada capítulo incluye código documentado y ejercicios prácticos que permiten al estudiante no solo comprender los conceptos teóricos, sino también aplicarlos inmediatamente en el análisis de datos reales. Esta integración de teoría y práctica computacional responde a las demandas del ejercicio profesional contemporáneo, donde el manejo de grandes volúmenes de datos es cada vez más frecuente. El libro también aborda consideraciones éticas fundamentales en la experimentación animal, integrando el principio de las 3Rs (Reemplazo, Reducción y Refinamiento) en el diseño experimental. Reconocemos que el uso de animales en investigación conlleva una responsabilidad moral que debe reflejarse en diseños estadísticamente sólidos que minimicen el número de animales utilizados mientras mantienen el poder estadístico necesario para obtener conclusiones válidas.

Los ejemplos y casos de estudio han sido cuidadosamente seleccionados para representar la diversidad de especies y contextos que caracterizan la medicina veterinaria moderna: desde animales de compañía hasta especies productivas, desde fauna silvestre hasta animales de laboratorio. Esta diversidad refleja la amplitud del campo veterinario y prepara al estudiante para enfrentar desafíos estadísticos en cualquier área de especialización que elija. Es importante destacar que este libro no pretende reemplazar la consulta con bioestadísticos especializados en proyectos de investigación complejos, sino proporcionar al médico veterinario las herramientas necesarias para comprender, interpretar y comunicar análisis estadísticos de manera competente. En una era donde la medicina basada en evidencia es el estándar de cuidado, la alfabetización estadística no es opcional sino esencial para el ejercicio profesional responsable.

Agradecemos a los numerosos colegas veterinarios, estadísticos y estudiantes que han contribuido con sus comentarios y sugerencias durante el desarrollo de esta obra. Sus aportes han sido invaluable para asegurar que el contenido sea tanto riguroso como accesible, manteniendo siempre la relevancia práctica que caracteriza a las ciencias veterinarias.

Esperamos que este libro se convierta en una herramienta fundamental en la formación de las nuevas generaciones de médicos veterinarios y zootecnistas, capacitándolos para contribuir al avance del conocimiento científico en beneficio de la salud animal, la salud pública y la producción sostenible de alimentos. La estadística, lejos de ser una disciplina árida y abstracta, es el lenguaje mediante el cual podemos

comprender la variabilidad biológica, evaluar intervenciones y, en última instancia, mejorar el bienestar de los animales bajo nuestro cuidado.

Lic. Víctor Alejandro Bosquez Barcenés PhD

ÍNDICE

PRÓLOGO.....	i
ÍNDICE.....	v
ÍNDICE DE TABLAS.....	xxiii
ÍNDICE DE FIGURAS	xxvii
INTRODUCCIÓN	xxxii
CAPÍTULO I.....	37
1 FUNDAMENTOS DE LA BIOESTADÍSTICA EN MEDICINA VETERINARIA	37
1.1 Introducción a la bioestadística y su importancia en medicina veterinaria.....	37
1.1.1 Objetivos del capítulo	38
1.2 Importancia de la estadística en medicina veterinaria.....	39
1.2.1 Fundamentos y conceptos básicos	39
1.2.2 Aplicaciones en la práctica clínica veterinaria	40
1.2.3 Aplicaciones en investigación veterinaria	42
1.2.4 Aplicaciones en producción animal y zootecnia.....	44
1.2.5 Aplicaciones en epidemiología y salud pública veterinaria.....	46

1.2.6	Estadística en la Era de Big Data veterinario	47
1.3	Tipos de variables en investigación veterinaria	48
1.3.1	Fundamentos conceptuales	48
1.3.2	Variables cuantitativas: Análisis detallado.....	48
1.3.3	Variables cualitativas: Clasificación comprehensiva	52
1.3.4	Variables según su papel en la investigación	56
1.4	Poblaciones y muestras en estudios animales	59
1.4.1	Conceptos fundamentales de población.....	59
1.4.2	Características poblacionales en contexto veterinario	61
1.4.3	Conceptos y teoría del muestreo	63
1.4.4	Cálculo del tamaño de muestra	68
1.4.5	Aspectos éticos del muestreo en animales.....	71
1.5	Escalas de medición	72
1.5.1	Fundamentos teóricos de las escalas de medición.....	72
1.5.2	Propiedades matemáticas de las escalas	73
1.5.3	Escala nominal en medicina veterinaria	73
1.5.4	Análisis estadístico de datos nominales.....	74

1.5.5	Escala ordinal en medicina veterinaria	75
1.5.6	Escala de intervalo en medicina veterinaria	78
1.5.7	Escala de razón en medicina veterinaria.....	80
1.5.8	Consideraciones prácticas y errores comunes	82
1.6	Organización y presentación de datos.....	83
1.6.1	Principios de organización de datos.....	83
1.6.2	Tablas de frecuencia y contingencia	87
1.6.3	Representación gráfica de datos	90
1.6.4	Medidas de resumen y estadísticos descriptivos.....	95
1.6.5	Presentación profesional de resultados	98
1.6.6	El método científico en investigación veterinaria.....	99
1.7	Conclusiones del capítulo	104
1.8	Autoevaluación del capítulo.....	105
CAPÍTULO II.....		107
2	ESTADÍSTICA DESCRIPTIVA	107
2.1	Objetivos de Aprendizaje	107
2.2	Introducción	107

2.3	Medidas de tendencia central	108
2.3.1	Conceptos fundamentales	108
2.3.2	Media aritmética	108
2.3.3	Mediana	110
2.3.4	Moda	111
2.3.5	Media ponderada.....	112
2.3.6	Implementación en R.....	113
2.4	Medidas de dispersión.....	114
2.4.1	Importancia de la variabilidad en medicina veterinaria.....	114
2.4.2	Rango.....	114
2.4.3	Varianza	115
2.4.4	Desviación estándar	115
2.4.5	Coeficiente de variación	117
2.4.6	Rango intercuartílico.....	118
2.5	Medidas de forma y posición.....	119
2.5.1	Asimetría (Skewness)	119
2.5.2	Curtosis	121

2.5.3	Percentiles y cuartiles	122
2.6	Representaciones gráficas	123
2.6.1	Principios de visualización de datos veterinarios	123
2.6.2	Histogramas	124
2.6.3	Diagramas de caja (Box Plots).....	125
2.6.4	Diagramas de dispersión.....	126
2.6.5	Gráficos de barras y columnas.....	127
2.6.6	Gráficos circulares (Pie Charts).....	128
2.6.7	Gráficos de líneas.....	129
2.6.8	Mapas de calor (Heatmaps)	130
2.7	Aplicaciones en parámetros clínicos veterinarios	131
2.7.1	Establecimiento de valores de referencia.....	131
2.7.2	Análisis de parámetros reproductivos.....	134
2.7.3	Monitoreo de crecimiento animal	136
2.7.4	Control de calidad en laboratorio veterinario	138
2.8	Conclusiones del capítulo	140
CAPÍTULO III		142

3	PROBABILIDAD Y DISTRIBUCIONES	142
3.1	Objetivos de aprendizaje	142
3.2	Conceptos básicos de probabilidad	142
3.2.1	Introducción a la probabilidad en ciencias veterinarias	142
3.2.2	Definiciones fundamentales.....	143
3.2.3	Enfoques de la probabilidad	143
3.2.4	Axiomas de Probabilidad.....	144
3.2.5	Reglas fundamentales de probabilidad	145
3.2.6	Teorema de bayes en medicina veterinaria	146
3.3	Distribución normal	147
3.3.1	Características y propiedades.....	147
3.3.2	Distribución normal estándar.....	148
3.3.3	Aplicaciones en medicina veterinaria	149
3.3.4	Verificación de normalidad.....	150
3.4	Distribución Binomial y de Poisson.....	150
3.4.1	Distribución Binomial.....	150
3.4.2	Distribución de Poisson	152

3.4.3	Comparación y selección de distribuciones.....	153
3.5	Distribuciones muestrales	153
3.5.1	Concepto de distribución muestral	153
3.5.2	Distribución muestral de la media	154
3.5.3	Distribución muestral de la proporción	155
3.5.4	Distribución t de Student	156
3.5.5	Distribución Chi-cuadrado (χ^2).....	156
3.6	Teorema del límite central.....	157
3.6.1	Enunciado y significado.....	157
3.6.2	Condiciones de aplicación	157
3.6.3	Implicaciones prácticas en veterinaria.....	158
3.6.4	Demostración práctica	158
3.6.5	Aplicaciones avanzadas	159
3.7	Problemas propuestos.....	160
3.8	Actividades prácticas con software estadístico	162
3.9	Conclusiones del capítulo	162
CAPÍTULO IV.....		165

4	INFERENCIA ESTADÍSTICA.....	165
4.1	Objetivos del capítulo	165
4.2	Introducción	165
4.3	Estimación puntual y por intervalos.....	166
4.3.1	Conceptos fundamentales	166
4.3.2	Propiedades de los buenos estimadores.....	167
4.3.3	Estimación puntual	167
4.3.4	Ejemplo práctico: Estimación del peso promedio en terneros	168
4.3.5	Distribución muestral del estimador	169
4.3.6	Estimación por intervalos	170
4.3.7	Ejemplo aplicado: Prevalencia de Brucelosis Bovina	170
4.4	Intervalos de confianza	171
4.4.1	Concepto e interpretación	171
4.4.2	Construcción de intervalos de confianza	172
4.4.3	Ejemplo práctico: Concentración de cortisol en perros con estrés	173
4.4.4	Intervalos de confianza para diferencias.....	175

4.4.5	Caso clínico: Comparación de tratamientos para mastitis bovina	175
4.4.6	Factores que afectan la amplitud del IC.....	177
4.5	Pruebas de hipótesis	179
4.5.1	Fundamentos conceptuales	179
4.5.2	Elementos de una prueba de hipótesis	179
4.5.3	Tipos de pruebas de hipótesis	180
4.5.4	Ejemplo práctico: Evaluación de un suplemento nutricional	181
4.5.5	Pruebas para diferentes situaciones	183
4.5.6	Relación entre intervalos de confianza y pruebas de hipótesis	184
4.6	Error tipo I y tipo II.....	185
4.6.1	Definición y consecuencias	185
4.6.2	Factores que afectan los errores.....	186
4.6.3	Ejemplo aplicado: Detección de residuos antibióticos	186
4.6.4	Control de errores en estudios veterinarios.....	188
4.6.5	Corrección por comparaciones múltiples	190
4.7	Potencia estadística	191

4.7.1	Concepto e importancia	191
4.7.2	Factores que determinan la potencia.....	192
4.7.3	Cálculo de potencia para diferentes diseños	192
4.7.4	Factores que determinan la potencia.....	193
4.7.5	Cálculo de potencia para diferentes diseños	193
4.7.6	Ejemplo práctico: Ensayo de eficacia vacunal	195
4.7.7	Análisis de potencia Post-Hoc	197
4.7.8	Potencia en diseños complejos	198
4.8	Tamaño de muestra	199
4.8.1	Importancia del cálculo a priori	199
4.8.2	Fórmulas básicas para tamaño de muestra.....	199
4.8.3	Aplicación práctica: Estudio de prevalencia.....	200
4.8.4	Tamaño de muestra para ensayos clínicos	202
4.8.5	Consideraciones especiales en veterinaria.....	204
4.9	Problemas Propuestos	210
4.10	Conclusiones del capítulo	211
CAPÍTULO V		214

5	PRUEBAS PARAMÉTRICAS.....	214
5.1	Objetivos del capítulo	214
5.2	Introducción	214
5.3	Prueba t de Student para una muestra	215
5.3.1	Fundamentos teóricos	215
5.3.2	Hipótesis	216
5.3.3	Aplicación en veterinaria: Caso práctico	216
5.3.4	Consideraciones prácticas.....	217
5.4	Prueba t para muestras independientes	217
5.4.1	Fundamentos teóricos	217
5.4.2	Prueba de Levene para homogeneidad de varianzas	218
5.5	Prueba t para muestras pareadas	218
5.5.1	Fundamentos teóricos	218
5.5.2	Aplicación en veterinaria: Caso práctico	219
5.5.3	Interpretación y relevancia clínica.....	220
5.5.4	Fundamentos teóricos	221
5.5.5	Aplicación en veterinaria: Caso práctico	221

5.5.6	Cálculos del ANOVA.....	222
5.6	ANOVA de dos factores	223
5.6.1	Fundamentos teóricos	223
5.6.2	Aplicación en veterinaria: Caso práctico	224
5.6.3	Interpretación de resultados	225
5.6.4	Gráfico de interacción.....	225
5.7	Pruebas Post-Hoc	226
5.7.1	Fundamentos y selección	226
5.7.2	Prueba de Tukey HSD.....	226
5.7.3	Aplicación práctica	226
5.7.4	Prueba de Dunnett.....	227
5.7.5	Corrección de Bonferroni	227
5.8	Supuestos y verificación	228
5.8.1	Supuestos de las pruebas paramétricas	228
5.8.2	Verificación de normalidad.....	228
5.8.3	Verificación de homogeneidad de varianzas.....	229
5.8.4	Verificación de independencia	230

5.8.5	Transformaciones de Datos.....	230
5.8.6	Análisis de residuos	231
5.8.7	Manejo de violaciones de supuestos.....	232
5.8.8	Robustez de las pruebas paramétricas	233
5.9	Ejercicios propuestos	233
5.10	Problemas resueltos.....	234
5.10.1	Problema 1: Evaluación de un tratamiento para artritis canina	234
5.10.2	Problema 2: Comparación de tres métodos de conservación de semen equino	236
5.10.3	Problema 3: Interacción entre raza y sistema de alimentación en ovinos.....	239
5.11	Actividades de Autoevaluación.....	241
5.12	Conclusiones	244
CAPÍTULO VI.....		247
6	PRUEBAS NO PARAMÉTRICAS EN CIENCIAS VETERINARIAS	247
6.1	Objetivos del capítulo	247
6.2	Introducción	247

6.3	Prueba de Mann-Whitney U.....	248
6.3.1	Fundamentos teóricos	248
6.3.2	Procedimiento de cálculo.....	248
6.3.3	Aplicación en medicina veterinaria	249
6.3.4	Implementación en Software Estadístico.....	250
6.4	Prueba de Wilcoxon para muestras pareadas	250
6.4.1	Fundamentos teóricos	250
6.4.2	Supuestos y condiciones	251
6.4.3	Procedimiento de cálculo.....	251
6.4.4	Aplicación práctica	252
6.5	Prueba de Kruskal-Wallis.....	253
6.5.1	Fundamentos teóricos	253
6.5.2	Estadístico de prueba	253
6.5.3	Aplicación en producción animal	254
6.5.4	Comparaciones Múltiples Post-hoc	255
6.6	Prueba de Friedman.....	256
6.6.1	Fundamentos teóricos	256

6.6.2	Diseño y aplicación.....	256
6.6.3	Caso veterinario: Evaluación de analgésicos.....	256
6.7	Pruebas de Chi-cuadrado.....	258
6.7.1	Fundamentos y aplicaciones	258
6.7.2	Cálculo del estadístico	258
6.7.3	Aplicación epidemiológica	258
6.7.4	Corrección de Yates	260
6.8	Prueba exacta de Fisher.....	260
6.8.1	Fundamentos teóricos	260
6.8.2	Aplicación en investigación clínica	260
6.8.3	Implementación computacional.....	261
6.9	Correlación de Spearman	262
6.9.1	Fundamentos teóricos	262
6.9.2	Cálculo y propiedades.....	262
6.9.3	Aplicación en medicina veterinaria	263
6.9.4	Comparación con correlación de Pearson.....	264
6.10	Actividades de autoevaluación.....	265

6.10.1 Ejercicios resueltos	265
6.11 Guía de Implementación en Software estadístico	267
6.11.1 Aplicación Integral con R.....	267
6.11.2 Análisis completo: Caso de estudio multivariado.....	268
6.11.3 Análisis de normalidad y decisión de prueba	270
6.11.4 Aplicación de Kruskal-Wallis con Post-hoc	271
6.11.5 Visualización profesional de resultados.....	272
6.12 Conclusiones	273
CAPÍTULO VII	277
7 ANÁLISIS DE CORRELACIÓN Y REGRESIÓN	277
7.1 Objetivos del Capítulo.....	277
7.2 Introducción	277
7.3 Correlación de Pearson	278
7.3.1 Fundamentos teóricos	278
7.3.2 Supuestos y condiciones de aplicación.....	279
7.3.3 Interpretación del coeficiente de correlación.....	279
7.3.4 Ejemplo práctico: Correlación entre peso corporal y frecuencia cardíaca en caninos	280

7.3.5	Prueba de significancia	282
7.4	Regresión lineal simple	282
7.4.1	Modelo de regresión lineal simple	282
7.4.2	Estimación de parámetros por mínimos cuadrados	283
7.4.3	Ejemplo aplicado: Predicción de glucemia según peso en felinos diabéticos	283
7.4.4	Coeficiente de determinación (R^2)	285
7.4.5	Intervalos de confianza y predicción	285
7.5	Regresión lineal múltiple	285
7.5.1	Modelo general	285
7.5.2	Estimación de parámetros	286
7.5.3	Caso práctico: Predicción de producción láctea en bovinos ..	286
7.5.4	Interpretación de coeficientes	287
7.5.5	Multicolinealidad	288
7.6	Regresión logística	288
7.6.1	Fundamentos del modelo logístico	288
7.7	Conclusiones	289
7.8	Anexos	293

7.8.1	Ejercicios Adicionales con Soluciones	293
7.8.2	Anexos técnicos	295
	GLOSARIO.....	299
	BIBLIOGRAFÍAS	302

ÍNDICE DE TABLAS

Tabla 1. Parámetros estadísticos para evaluación de pruebas diagnósticas veterinarias	41
Tabla 2. Aplicaciones estadísticas en mejoramiento genético animal.	45
Tabla 3. Comparación entre variables continuas y discretas en contexto veterinario.	51
Tabla 4. Escalas ordinales validadas en medicina veterinaria.	55
Tabla 5. Consideraciones especiales según tipo de población animal.	62
Tabla 6. Estudio de mastitis en 10,000 vacas lecheras estratificadas por número de lactancia.....	65
Tabla 7. Tamaños de muestra para diferentes escenarios veterinarios.	70
Tabla 8. Aplicaciones de escalas nominales en diferentes áreas veterinarias.	75
Tabla 9. Sistema IRIS para Enfermedad Renal Crónica.	77
Tabla 10. Escalas de intervalo y cuasi-intervalo en veterinaria.	79
Tabla 11. Evaluación de crecimiento en corderos con diferentes dietas.	81

Tabla 12. Guía de selección de análisis según escala de medición....	83
Tabla 13. Protocolo de validación de datos veterinarios.	86
Tabla 14. Distribución de diagnósticos en 500 consultas de emergencia veterinaria.....	87
Tabla 15. Relación entre método diagnóstico y resultado en detección de Brucella en 300 bovinos.	88
Tabla 16. Distribución de pesos al sacrificio de 200 cerdos.	89
Tabla 17. Selección de gráficos según tipo de datos y objetivo.	93
Tabla 18. Medidas de dispersión según escala y distribución.	96
Tabla 19. Registro mensual de casos de urgencia por categoría diagnóstica.....	111
Tabla 20. Ganancia de peso en lotes bovinos.	112
Tabla 21. Análisis de dos parámetros en bovinos.	117
Tabla 22. Comparación y selección de distribuciones.....	153
Tabla 23. Tabla de decisiones y errores.	186
Tabla 24. Datos (g/dL).	219
Tabla 25. La tabla ANOVA.	221
Tabla 26. Datos resumidos.....	222

Tabla 27. Tabla ANOVA.	222
Tabla 28. Datos resumidos.	224
Tabla 29. Tabla ANOVA de dos factores.	225
Tabla 30. Comparaciones pareadas.	227
Tabla 31. Datos.	235
Tabla 32. Datos.	236
Tabla 33. Tabla ANOVA.	238
Tabla 34. Pruebas post-hoc (Tukey).	238
Tabla 35. Datos resumidos.	239
Tabla 36. Tabla ANOVA factorial.	240
Tabla 37. Reducción en HPG según tratamiento.	249
Tabla 38. Niveles de glucemia pre y post-tratamiento.	252
Tabla 39. GDP según tipo de dieta.	254
Tabla 40. Scores de dolor según protocolo analgésico.	257
Tabla 41. Tabla de contingencia 2×2.	259
Tabla 42. Respuesta al tratamiento.	260
Tabla 43. Datos de SCC e intervalo parto-concepción.	263

Tabla 44. Interpretación del Coeficiente de Correlación de Pearson.	279
Tabla 45. Datos de Peso y Glucemia en Felinos Diabéticos.....	283
Tabla 46. Análisis de varianza del modelo de regresión múltiple. ..	287
Tabla 47. Valores críticos comunes.....	296
Tabla 48. Guía rápida de selección de medidas descriptivas.	298

ÍNDICE DE FIGURAS

Figura 1. Ejemplo de estructura correcta.....	84
Figura 2. Verificación de rangos.....	85
Figura 3. Análisis de Concordancia.....	89
Figura 4 . Ejemplo Práctico	91
Figura 5. Ejemplo Código R.....	92
Figura 6. Comparación de glucemia en tres estados metabólicos de vacas lecheras.....	92
Figura 7. Supervivencia de perros con linfoma según protocolo de quimioterapia.....	94
Figura 8. Ejemplo de Código R para análisis veterinario.....	99
Figura 9. Fórmula.	112
Figura 10. Implementación en R.	113
Figura 11. Fórmula.	114
Figura 12. Fórmula (muestra).....	115
Figura 13. Fórmula.	115
Figura 14. Análisis de cortisol en perros con estrés.	116
Figura 15. Fórmula.	117

Figura 16. Fórmula.	118
Figura 17. Análisis de creatinina en gatos geriátricos.	119
Figura 18. Distribución de edad en primera gestación en bovinos..	120
Figura 19. Análisis de proteínas totales en aves.	122
Figura 20. Curva de crecimiento en lechones.	123
Figura 21. Distribución de pesos en cerdos de engorde.	125
Figura 22. Comparación de ALT entre razas caninas.	126
Figura 23. Relación peso-edad en potros.	127
Figura 24. Prevalencia de enfermedades en una clínica veterinaria.	128
Figura 25. Distribución de especies atendidas.	129
Figura 26. Evolución de producción láctea.	130
Figura 27. Correlación entre parámetros hematológicos.....	131
Figura 28. Valores de referencia para glucosa en felinos.	133
Figura 29. Análisis descriptivo de intervalos entre partos en bovinos.	135
Figura 30. Curvas de crecimiento en pollos de engorde.	137
Figura 31. Gráficos de control de Levey-Jennings.....	139

Figura 32. Práctica en R	162
Figura 33. Estimación del peso promedio en terneros.	169
Figura 34. Prevalencia de Brucelosis Bovina.....	171
Figura 35. Concentración de cortisol en perros con estrés	174
Figura 36. Comparación de tratamientos para mastitis bovina.	176
Figura 37. Factores que afectan la amplitud del IC.....	178
Figura 38. Evaluación de un suplemento nutricional.	182
Figura 39. Prueba t para dos muestras independientes.....	183
Figura 40. Prueba t pareada.....	184
Figura 41. Prueba de proporciones.....	184
Figura 42. Relación entre intervalos de confianza y pruebas de hipótesis.....	185
Figura 43. Detección de residuos antibióticos.....	187
Figura 44. Control de errores en estudios veterinarios.....	189
Figura 45. Corrección por comparaciones múltiples.....	191
Figura 46. Potencia para prueba t de una muestra.....	194
Figura 47. Ensayo de eficacia vacunal.	196
Figura 48. Análisis de potencia Post-Hoc.	197

Figura 49. Potencia en diseños complejos.....	198
Figura 50. Estudio de prevalencia.	201
Figura 51. Tamaño de muestra para ensayos clínicos.	203
Figura 52. Efecto de Diseño (DEFF) en muestreo por conglomerados.	204
Figura 53. Método de ecuación de recursos.	205
Figura 54. Comparación de Tratamientos.	207
Figura 55. Cálculo de tamaño de muestra.	209
Figura 56. Cálculo de tamaño de muestra.	210
Figura 57. Ejemplo de aplicación en R.	229
Figura 58. Ejemplo de transformación logarítmica.....	231
Figura 59. Verificación de supuestos.....	237
Figura 60. Ejercicio con R.....	243
Figura 61. Código en R.	250
Figura 62. Implementación en R.	255
Figura 63. Implementación computacional.	261
Figura 64. Ejemplo comparativo con software.....	265
Figura 65. Configuración del entorno de trabajo.....	268

Figura 66. Caso de estudio multivariado.....	269
Figura 67. Análisis de normalidad y decisión de prueba.....	270
Figura 68. Aplicación de Kruskal-Wallis con Post-hoc.	271
Figura 69. Visualización profesional de resultados.	272
Figura 70. Datos recolectados (n=20).	281
Figura 71. Código R para análisis básicos.....	297

INTRODUCCIÓN

La bioestadística constituye un pilar fundamental en el desarrollo de las ciencias veterinarias modernas, proporcionando el marco metodológico necesario para transformar observaciones empíricas en conocimiento científico validado y reproducible. En un contexto donde la medicina veterinaria enfrenta desafíos cada vez más complejos —desde el control de enfermedades emergentes hasta la optimización de sistemas productivos sostenibles— la capacidad de analizar e interpretar datos de manera rigurosa se ha convertido en una competencia esencial para el profesional veterinario del siglo XXI.

El presente libro, "Estadística y Diseño Experimental en Ciencias Veterinarias", surge como respuesta a la necesidad imperante de proporcionar a los estudiantes de Medicina Veterinaria y Zootecnia una obra integral que aborde los métodos estadísticos desde una perspectiva genuinamente aplicada al contexto veterinario. A diferencia de los textos genéricos de bioestadística que tradicionalmente han dominado la bibliografía académica, esta obra ha sido concebida específicamente para abordar las particularidades y desafíos únicos que caracterizan la investigación y práctica en ciencias veterinarias.

La relevancia de la estadística en medicina veterinaria trasciende el ámbito académico para convertirse en una herramienta indispensable en la práctica profesional cotidiana. En el contexto clínico, el veterinario debe interpretar constantemente resultados de pruebas diagnósticas, establecer intervalos de referencia específicos para diferentes especies y razas, evaluar la eficacia de tratamientos y tomar decisiones

terapéuticas basadas en evidencia cuantificable. En el ámbito de la producción animal, la optimización de parámetros productivos, el mejoramiento genético y la evaluación económica de sistemas de producción dependen fundamentalmente del análisis estadístico riguroso. Asimismo, en salud pública veterinaria, la vigilancia epidemiológica, el control de zoonosis y la evaluación de riesgos sanitarios requieren un dominio profundo de métodos estadísticos aplicados.

La estructura pedagógica de este libro ha sido cuidadosamente diseñada para facilitar un aprendizaje progresivo y significativo. Los primeros capítulos establecen los fundamentos de la bioestadística y la estadística descriptiva, proporcionando las bases conceptuales necesarias para comprender análisis más complejos. Los capítulos intermedios abordan la teoría de probabilidades, la inferencia estadística y las pruebas de hipótesis, tanto paramétricas como no paramétricas, siempre contextualizadas mediante ejemplos extraídos de situaciones reales en medicina veterinaria. Los capítulos finales se adentran en el diseño experimental y técnicas estadísticas avanzadas, preparando al estudiante para la investigación científica rigurosa y la toma de decisiones basadas en evidencia.

Una característica distintiva de esta obra es su enfoque práctico y aplicado. Cada concepto teórico se acompaña de ejemplos específicos del ámbito veterinario, casos de estudio detallados y ejercicios que reflejan situaciones reales que el futuro profesional encontrará en su práctica. Desde el análisis de parámetros hematológicos en pequeños animales hasta la evaluación de sistemas de producción en especies

mayores, los ejemplos abarcan la diversidad de contextos que caracterizan las ciencias veterinarias contemporáneas.

La integración de herramientas computacionales constituye otro aspecto fundamental de este libro. Reconociendo que el análisis estadístico moderno depende crucialmente del uso de software especializado, se incluyen guías detalladas para la implementación de análisis utilizando R, un lenguaje de programación estadística de código abierto que se ha convertido en el estándar de facto en investigación biomédica. Cada capítulo incluye código documentado y ejercicios prácticos que permiten al estudiante no solo comprender los conceptos teóricos, sino también aplicarlos inmediatamente en el análisis de datos reales.

El libro también aborda consideraciones éticas fundamentales en la experimentación animal, integrando el principio de las 3Rs (Reemplazo, Reducción y Refinamiento) en el diseño experimental. Se enfatiza que el uso responsable de animales en investigación requiere diseños estadísticamente sólidos que minimicen el número de individuos utilizados mientras mantienen el poder estadístico necesario para obtener conclusiones válidas. Esta perspectiva ética se entrelaza con el rigor metodológico a lo largo de toda la obra.

Los ejemplos y casos de estudio han sido cuidadosamente seleccionados para representar la diversidad de especies y contextos que caracterizan la medicina veterinaria moderna: desde animales de compañía hasta especies productivas, desde fauna silvestre hasta animales de laboratorio. Esta diversidad refleja la amplitud del campo veterinario y

prepara al estudiante para enfrentar desafíos estadísticos en cualquier área de especialización que elija desarrollar en su carrera profesional.

Es importante destacar que este libro no pretende reemplazar la consulta con bioestadísticos especializados en proyectos de investigación particularmente complejos, sino proporcionar al médico veterinario las herramientas necesarias para comprender, interpretar y comunicar análisis estadísticos de manera competente. En una era donde la medicina basada en evidencia constituye el estándar de cuidado, la alfabetización estadística no es opcional sino esencial para el ejercicio profesional responsable y ético.

La obra se estructura en siete capítulos principales que abordan progresivamente desde los conceptos fundamentales hasta técnicas avanzadas de análisis. Esta organización permite tanto un estudio secuencial para estudiantes de pregrado como una consulta específica para profesionales en ejercicio que requieran profundizar en técnicas particulares.

Esperamos que este libro se convierta en una herramienta fundamental en la formación de las nuevas generaciones de médicos veterinarios y zootecnistas, capacitándolos para contribuir al avance del conocimiento científico en beneficio de la salud animal, la salud pública y la producción sostenible de alimentos. La estadística, lejos de ser una disciplina árida y abstracta, es el lenguaje mediante el cual podemos comprender la variabilidad biológica inherente a los sistemas vivos, evaluar intervenciones terapéuticas y preventivas, y en última instancia,

mejorar el bienestar de los animales bajo nuestro cuidado y la calidad de vida de las sociedades que dependen de ellos.

CAPÍTULO I

1 FUNDAMENTOS DE LA BIOESTADÍSTICA EN MEDICINA VETERINARIA

1.1 Introducción a la bioestadística y su importancia en medicina veterinaria

La bioestadística constituye una disciplina científica que integra métodos estadísticos con la biología y la medicina, permitiendo analizar, interpretar y comunicar información derivada de fenómenos vitales. Su aplicación en medicina veterinaria resulta esencial, ya que los animales, al igual que los seres humanos, presentan variabilidad biológica, condiciones ambientales diversas y respuestas heterogéneas a tratamientos clínicos o protocolos productivos. En este sentido, la bioestadística se convierte en un pilar metodológico para garantizar la validez de los hallazgos y la confiabilidad de las decisiones en salud y producción animal (Ambarcıoğlu, 2021).

El papel de la bioestadística en veterinaria es amplio y abarca desde el diseño de experimentos en estudios de nutrición, reproducción o farmacología, hasta la evaluación de pruebas diagnósticas y la vigilancia epidemiológica de enfermedades animales. A través de medidas descriptivas, pruebas de hipótesis, modelos de regresión y técnicas multivariantes, se logra transformar datos en conocimiento útil para la práctica clínica, la gestión de granjas y la investigación científica (Costa Silva et al., 2024).

En el ámbito académico, la enseñanza de la bioestadística fomenta en los futuros profesionales el desarrollo del pensamiento crítico y la capacidad de interpretar literatura científica. Además, proporciona las herramientas necesarias para elaborar protocolos de investigación sólidos, analizar resultados con rigor y comunicar hallazgos bajo estándares internacionales. En un mundo donde la producción animal debe ser sostenible y respetuosa con el bienestar, el uso adecuado de métodos estadísticos se convierte en una garantía de calidad y transparencia.

La importancia de la bioestadística trasciende lo técnico, pues también se vincula con la ética y la responsabilidad social. Un análisis inadecuado puede conducir a conclusiones erróneas que afecten tanto la salud animal como la confianza de productores, tutores y consumidores. Por ello, la integración de herramientas estadísticas en la práctica veterinaria no es opcional, sino un requisito para una profesión que busca ser científica, innovadora y socialmente comprometida (Mata, 2022).

1.1.1 Objetivos del capítulo

Al finalizar este capítulo, el estudiante será capaz de:

- Comprender la importancia y aplicaciones de la estadística en el ejercicio profesional veterinario
- Identificar y clasificar correctamente los tipos de variables utilizadas en investigación veterinaria

- Diferenciar entre poblaciones y muestras, aplicando criterios de selección apropiados para estudios animales
- Reconocer y aplicar las escalas de medición en el contexto veterinario
- Organizar, presentar e interpretar datos mediante técnicas estadísticas básicas

1.2 Importancia de la estadística en medicina veterinaria

1.2.1 Fundamentos y conceptos básicos

La bioestadística constituye una disciplina fundamental en las ciencias veterinarias, proporcionando el marco metodológico necesario para la recolección, organización, análisis e interpretación de datos relacionados con la salud animal, la producción pecuaria y la salud pública veterinaria. Su importancia radica en la capacidad de transformar observaciones empíricas en conocimiento científico válido y reproducible, permitiendo al médico veterinario zootecnista tomar decisiones basadas en evidencia cuantitativa más que en impresiones subjetivas (Lušničenko et al., 2024).

En el contexto de la medicina veterinaria moderna, donde coexisten múltiples especies animales con características fisiológicas diversas, sistemas de producción variados y desafíos sanitarios complejos, la estadística proporciona las herramientas necesarias para manejar la variabilidad inherente a los sistemas biológicos. Esta variabilidad, lejos de ser un obstáculo, constituye información valiosa que, correctamente analizada, permite comprender los procesos de salud y enfermedad,

optimizar tratamientos y mejorar los sistemas de producción animal (Kumar & Yadav, 2021).

La evolución histórica de la bioestadística veterinaria se remonta a los primeros estudios epidemiológicos en poblaciones animales durante el siglo XIX, cuando pioneros como John Snow aplicaron principios estadísticos básicos para comprender la transmisión de enfermedades. En la actualidad, con el advenimiento de tecnologías de análisis masivo de datos y la medicina veterinaria de precisión, la estadística se ha convertido en una competencia esencial para el profesional veterinario del siglo XXI.

1.2.2 Aplicaciones en la práctica clínica veterinaria

a) Diagnóstico y valores de referencia

La práctica clínica diaria requiere constantemente la interpretación de datos numéricos provenientes de análisis laboratoriales, estudios imagenológicos y exámenes físicos. La estadística proporciona los métodos para establecer intervalos de referencia específicos para cada especie, raza, edad y condición fisiológica, permitiendo distinguir entre variaciones normales y hallazgos patológicos significativos (Weng & Messam, 2025).

b) Evaluación de pruebas diagnósticas

La estadística permite evaluar la precisión diagnóstica de las pruebas mediante el cálculo de sensibilidad, especificidad, valores predictivos y

likelihood ratios. Estos parámetros son fundamentales para la interpretación adecuada de resultados y la toma de decisiones clínicas.

Tabla 1. Parámetros estadísticos para evaluación de pruebas diagnósticas veterinarias

Parámetro	Definición	Fórmula	Interpretación Clínica	Ejemplo Veterinario
Sensibilidad	Probabilidad de resultado positivo en animales enfermos	$VP/(VP+FN)$	Alta sensibilidad: ideal para screening	Test ELISA para parvovirus: 95%
Especificidad	Probabilidad de resultado negativo en animales sanos	$VN/(VN+FP)$	Alta especificidad: ideal para confirmación	PCR para leptospirosis: 98%
VPP	Probabilidad de enfermedad ante resultado positivo	$VP/(VP+FP)$	Depende de la prevalencia	Variable según población
VPN	Probabilidad de salud ante resultado negativo	$VN/(VN+FN)$	Útil para descartar enfermedad	Test rápido FeLV: 99%

LR+	Razón de Sens/(1-Esp) verosimilitud positiva	>10: cambio grande en probabilidad	Radiografía para displasia: 8.5
LR-	Razón de (1-Sens)/Esp verosimilitud negativa	<0.1: descarta enfermedad	Ecografía cardíaca: 0.05

Nota. VP: Verdaderos Positivos, VN: Verdaderos Negativos, FP: Falsos Positivos, FN: Falsos Negativos.

c) Monitoreo terapéutico

El seguimiento de la respuesta al tratamiento requiere análisis estadísticos para determinar si los cambios observados son clínicamente significativos o simplemente variaciones aleatorias. Las pruebas de hipótesis pareadas y el análisis de medidas repetidas permiten evaluar la eficacia de intervenciones terapéuticas de manera objetiva.

1.2.3 Aplicaciones en investigación veterinaria

a) Tamaño de muestra

La investigación veterinaria requiere diseños experimentales rigurosos que consideren las particularidades de trabajar con animales. El cálculo del tamaño de muestra debe equilibrar el poder estadístico necesario con consideraciones éticas sobre el uso de animales en experimentación, siguiendo el principio de las 3R (Reemplazo, Reducción y Refinamiento).

Ejemplo Práctico Detallado: Un grupo de investigación planea evaluar la eficacia de una nueva vacuna contra la bronquitis infecciosa aviar. Utilizando la fórmula para comparación de proporciones:

$$n = 2 \times [(Z_{1-\alpha/2} + Z_{1-\beta})^2 \times p \times (1-p)] / (p_1 - p_2)^2$$

Donde:

- $\alpha = 0.05$ (error tipo I)
- $\beta = 0.20$ (error tipo II, poder = 80%)
- $p_1 = 0.85$ (eficacia esperada de la nueva vacuna)
- $p_2 = 0.65$ (eficacia de la vacuna estándar)
- $p = (p_1 + p_2)/2 = 0.75$

El cálculo indica que se requieren 62 aves por grupo. Considerando una mortalidad esperada del 10%, el estudio debe iniciarse con 69 aves por grupo, totalizando 138 animales. Este cálculo garantiza poder estadístico adecuado mientras minimiza el número de animales utilizados (Mishra & Khan, 2024).

b) Análisis de Datos Complejos

La investigación veterinaria moderna genera conjuntos de datos cada vez más complejos que requieren técnicas estadísticas avanzadas:

1. **Datos Longitudinales:** Estudios de crecimiento, progresión de enfermedades crónicas y evaluación de tratamientos a largo plazo requieren modelos mixtos que consideren la correlación entre mediciones repetidas en el mismo animal.

2. **Datos Jerárquicos:** En estudios de campo, los animales están anidados dentro de corrales, granjas o regiones, requiriendo modelos multinivel para manejar adecuadamente la estructura de los datos.
3. **Datos de Supervivencia:** El análisis de tiempo hasta un evento (muerte, recidiva, destete) utiliza métodos especializados como las curvas de Kaplan-Meier y modelos de riesgos proporcionales de Cox.

1.2.4 Aplicaciones en producción animal y zootecnia

a) Evaluación de parámetros productivos

La optimización de sistemas de producción animal depende del análisis estadístico riguroso de indicadores productivos. Los parámetros como ganancia diaria de peso, conversión alimenticia, producción láctea y eficiencia reproductiva requieren monitoreo continuo y análisis estadístico para identificar oportunidades de mejora.

Ejemplo Práctico Detallado: En una granja porcina con 500 cerdas reproductoras, se implementó un sistema de análisis estadístico para monitorear la productividad. Los datos recolectados durante un año mostraron:

- Promedio de lechones nacidos vivos: 12.8 ± 2.3 (media $\pm$ DE)
- Mortalidad pre-destete: 11.5% (IC 95%: 10.8-12.2%)
- Peso al destete (21 días): 6.2 ± 0.8 kg
- Lechones destetados por cerda/año: 26.4

El análisis de regresión múltiple identificó que la paridad de la cerda ($p<0.001$), la estación del año ($p=0.003$) y el semental utilizado ($p=0.012$) eran factores significativos que afectaban el número de lechones destetados. Basándose en estos resultados, se implementaron estrategias diferenciadas de manejo que incrementaron la productividad en 8% el año siguiente (Leaverton & Zhu, 2025).

b) Mejoramiento genético

La selección genética en animales domésticos utiliza modelos estadísticos complejos para estimar valores genéticos y predecir el progreso genético. Los modelos animales BLUP (Best Linear Unbiased Prediction) y más recientemente, la selección genómica, requieren conocimientos profundos de estadística para su implementación e interpretación.

Tabla 2. Aplicaciones estadísticas en mejoramiento genético animal.

Componente	Método Estadístico	Aplicación Práctica	Software Especializado
Heredabilidad	Análisis de varianza, Modelos mixtos	Estimar proporción de un carácter genética	WOMBAT, ASReml
Correlaciones genéticas	Análisis multivariado	Predicción de respuesta correlacionada	DMU, BLUPF90
Valores genéticos	BLUP, ssGBLUP	Rankings para selección	PEST, MiX99

Consanguinidad	Coeficientes de parentesco	Control de endogamia	de ENDOG, CFC
Selección genómica	Modelos bayesianos	Predicción con marcadores SNP	BGLR, BayesR

1.2.5 Aplicaciones en epidemiología y salud pública veterinaria

a) Vigilancia epidemiológica

Los sistemas de vigilancia epidemiológica dependen fundamentalmente del análisis estadístico para detectar cambios en los patrones de enfermedad, identificar brotes y evaluar tendencias temporales. Las técnicas de series temporales, análisis espacial y modelos predictivos son herramientas esenciales en la epidemiología veterinaria moderna.

Ejemplo Práctico Detallado: Durante un brote de influenza aviar H5N1 en una región productora de aves, se implementó un sistema de vigilancia activa. El análisis estadístico espacial mediante la función K de Ripley identificó clusters significativos de casos en un radio de 3 km alrededor de granjas índice ($p < 0.001$). Los modelos de regresión logística identificaron los siguientes factores de riesgo:

- Distancia a cuerpos de agua (OR=3.2, IC 95%: 2.1-4.8)
- Densidad de granjas en 5 km (OR=1.8 por cada 10 granjas adicionales)
- Bioseguridad deficiente (OR=5.7, IC 95%: 3.9-8.3)

Esta información permitió focalizar las medidas de control, logrando contener el brote en 45 días con pérdidas 60% menores a las proyectadas inicialmente (Torres & Normando, 2021).

b) Control de Zoonosis

Las enfermedades zoonóticas representan un desafío mayor en salud pública. El análisis estadístico permite evaluar la dinámica de transmisión entre animales y humanos, identificar reservorios y evaluar estrategias de control.

1.2.6 Estadística en la Era de Big Data veterinario

a) Análisis de Datos Masivos

La digitalización de registros veterinarios, el uso de sensores en producción animal y la secuenciación genómica generan volúmenes masivos de datos que requieren nuevos enfoques estadísticos:

1. **Machine Learning:** Algoritmos de aprendizaje automático para predicción de enfermedades y optimización productiva
2. **Deep Learning:** Análisis de imágenes diagnósticas y reconocimiento de patrones complejos
3. **Data Mining:** Descubrimiento de patrones en bases de datos veterinarias grandes

Ejemplo de Implementación: Un hospital veterinario implementó un sistema de machine learning para predecir el riesgo de readmisión en pacientes caninos post-quirúrgicos. Utilizando datos de 5,000 casos históricos y 47 variables predictoras, el modelo Random Forest alcanzó

una precisión del 82% ($AUC = 0.89$), permitiendo identificar pacientes de alto riesgo para seguimiento intensivo (Sarma et al., 2022).

1.3 Tipos de variables en investigación veterinaria

1.3.1 Fundamentos conceptuales

Una variable constituye cualquier característica, atributo o propiedad observable que puede variar entre individuos de una población. En el contexto veterinario, las variables abarcan desde mediciones fisiológicas básicas hasta complejos indicadores de bienestar animal o productividad. La correcta identificación y clasificación de variables determina no solo el tipo de análisis estadístico aplicable, sino también la validez de las conclusiones derivadas del estudio (Karabulut, 2022).

La naturaleza de las variables en investigación veterinaria presenta particularidades únicas derivadas de la diversidad de especies estudiadas, los diferentes sistemas de producción y las múltiples dimensiones de la salud animal. A diferencia de la medicina humana, donde el foco se centra en una sola especie, la medicina veterinaria debe considerar las variaciones inter e intraespecíficas, lo que añade complejidad al proceso de definición y medición de variables.

1.3.2 Variables cuantitativas: Análisis detallado

a) Variables cuantitativas continuas

Las variables continuas pueden tomar cualquier valor dentro de un rango determinado, limitadas únicamente por la precisión del

instrumento de medición. En medicina veterinaria, constituyen la base de múltiples evaluaciones clínicas y productivas.

Características principales:

- Admiten valores decimales infinitos dentro de su rango
- Permiten operaciones aritméticas completas
- Generalmente siguen distribuciones conocidas (normal, log-normal, gamma)
- Requieren instrumentos de medición calibrados

Ejemplos veterinarios específicos:

1. Variables fisiológicas:

- Temperatura corporal: En bovinos oscila entre 37.5-39.5°C, con variaciones decimales importantes para el diagnóstico
- Frecuencia cardíaca: Variable según especie (equinos: 28-44 lpm; caninos: 70-120 lpm)
- Presión arterial: Medida en mmHg con variaciones continuas durante el ciclo cardíaco

2. Variables bioquímicas:

- Glucemia: Valores continuos expresados en mg/dL o mmol/L
- Proteínas séricas: Medidas con precisión de décimas de g/dL
- Hormonas: Concentraciones en pg/mL, ng/mL o UI/mL según el analito

3. Variables morfométricas:

- Longitud corporal: Mediciones en centímetros con decimales
- Circunferencia torácica: Utilizada para estimación de peso vivo
- Altura a la cruz: Importante en evaluación de conformación

Ejemplo práctico detallado: En un estudio de farmacocinética de enrofloxacin en alpacas, se midieron las concentraciones plasmáticas del fármaco a diferentes tiempos post-administración. Los valores obtenidos variaron continuamente desde 0.12 µg/mL a los 15 minutos hasta concentraciones indetectables (<0.01 µg/mL) a las 48 horas. El análisis de estos datos continuos mediante modelos no compartimentales permitió calcular:

- Concentración máxima (C_{max}): 2.84 ± 0.36 µg/mL
- Tiempo de vida media (t_{1/2}): 7.2 ± 1.1 horas
- Área bajo la curva (AUC): 18.4 ± 2.7 µg·h/mL

Estos parámetros farmacocinéticos, todos variables continuas, son esenciales para establecer regímenes de dosificación apropiados (Piper et al., 2023).

b) Variables cuantitativas discretas

Las variables discretas toman valores enteros específicos, generalmente resultantes de procesos de conteo. Aunque numéricas, no admiten valores intermedios entre dos valores consecutivos.

Características principales:

- Solo pueden tomar valores enteros
- Resultado de conteos o enumeraciones

- Frecuentemente siguen distribuciones discretas (Poisson, binomial negativa)
- No tienen sentido los valores fraccionarios

Ejemplos veterinarios específicos:

1. Variables reproductivas:

- Tamaño de camada: Número de crías por parto
- Número de servicios por concepción
- Días abiertos (aunque medido en días, es discreto)

2. Variables parasitológicas:

- Conteo de huevos por gramo de heces (HPG)
- Número de ectoparásitos por animal
- Carga de ooquistes en muestras fecales

3. Variables celulares:

- Recuento de células somáticas en leche
- Conteo diferencial de leucocitos
- Número de espermatozoides por mL

Tabla 3. Comparación entre variables continuas y discretas en contexto veterinario.

Aspecto	Variables Continuas	Variables Discretas	Implicación Práctica
Valores posibles	Infinitos en un rango	Finitos contables	o Afecta la precisión de medición

Instrumento de medición	de Escalas analógicas	Conteo directo	Determina el protocolo de medición
Distribución típica	Normal, log-normal	Poisson, binomial	Define el análisis estadístico
Ejemplo clínico	Creatinina sérica	Recuento plaquetario	Interpretación diferente
Transformación común	Logarítmica	Raíz cuadrada	Para normalizar datos
Error de medición	Error continuo	Error de conteo	Afecta intervalos de confianza

1.3.3 Variables cualitativas: Clasificación comprehensiva

a) Variables cualitativas nominales

Las variables nominales representan categorías sin orden intrínseco. La asignación de números a estas categorías es arbitraria y solo sirve para codificación, no implicando ninguna relación matemática.

Características fundamentales:

- No existe jerarquía entre categorías
- La codificación numérica es arbitraria
- Solo admiten análisis de frecuencias y proporciones
- Requieren pruebas estadísticas no paramétricas específicas

Clasificación detallada en veterinaria:

1. Variables taxonómicas:

- Especie: Bos taurus, Equus caballus, Canis familiaris
- Raza: Más de 400 razas caninas reconocidas por la FCI
- Familia/Orden: Clasificación zoológica

2. Variables clínicas:

- Diagnóstico principal: Categorías CIE-Vet
- Tipo de cirugía realizada
- Método de anestesia empleado

3. Variables de manejo:

- Sistema de producción: Intensivo, extensivo, semi-intensivo
- Tipo de instalaciones: Jaula, corral, pastoreo
- Protocolo vacunal aplicado

Ejemplo práctico detallado: En un estudio multicéntrico sobre dermatitis atópica canina, se clasificaron 500 casos según:

- Raza afectada (35 razas diferentes)
- Alérgeno identificado (ácaros, polen, alimentos, múltiple)
- Tratamiento inicial (corticosteroides, ciclosporina, oclacitinib, lokivetmab)

El análisis mediante tablas de contingencia y prueba Chi-cuadrado reveló asociación significativa entre raza y tipo de alérgeno ($\chi^2 = 67.3$,

$p < 0.001$), con los West Highland White Terriers mostrando mayor prevalencia de sensibilización a ácaros del polvo (72% vs 45% en la población general). Estos hallazgos, aunque las variables son nominales, tienen implicaciones importantes para el abordaje diagnóstico según la raza (Enoyoze et al., 2024).

b) Variables cualitativas ordinales

Las variables ordinales presentan un orden o jerarquía natural entre sus categorías, aunque las distancias entre categorías consecutivas no necesariamente son iguales.

Características fundamentales:

- Existe un orden lógico entre categorías
- Las distancias entre categorías pueden no ser uniformes
- Permiten análisis de tendencias
- Admiten pruebas estadísticas específicas para datos ordinales

Aplicaciones veterinarias específicas:

1. Escalas de evaluación clínica:

- Condición corporal (BCS): Escala de 1-9 en equinos y bovinos, 1-5 en pequeños animales
- Grado de deshidratación: Leve (<5%), Moderada (5-10%), Severa (>10%)
- Score de dolor: Escalas validadas como Glasgow Composite Pain Scale

2. Clasificación de severidad:

- Grado de displasia de cadera: A, B, C, D, E según FCI
- Estadiaje tumoral: Sistema TNM adaptado a veterinaria
- Clasificación de soplos cardíacos: Grados I-VI

3. Escalas productivas:

- Marmoreo en carne: Escala USDA adaptada
- Calidad de semen: Excelente, Bueno, Regular, Pobre
- Estado de madurez de forrajes: Vegetativo, Prefloración, Floración, Semillado

Tabla 4. Escalas ordinales validadas en medicina veterinaria.

Escala	Especie	Rango	Aplicación	Validación
BCS Laflamme	Caninos/Felinos	1-9	Evaluación nutricional	Correlación con % grasa ($r=0.87$)
Glasgow Pain Scale	Caninos	0-24	Dolor post- quirúrgico	Sensibilidad 92%, Especificidad 88%
Escala IRIS	Todas	1-4	Enfermedad renal crónica	Correlación con supervivencia
Score APGAR	Neonatos	0-10	Viabilidad neonatal	Predictor mortalidad <24h
Escala Obel	Equinos	1-4	Laminitis	Concordancia inter- observador $\kappa=0.78$

1.3.4 Variables según su papel en la investigación

a) Variables independientes

La variable independiente representa el factor que el investigador manipula o selecciona para estudiar su efecto. En investigación veterinaria experimental, estas variables son controladas deliberadamente, mientras que en estudios observacionales son características preexistentes que se miden.

Tipos en investigación veterinaria:

1. Variables de tratamiento:

- Dosis de fármaco administrada
- Tipo de vacuna aplicada
- Protocolo quirúrgico empleado

2. Variables ambientales:

- Temperatura ambiental
- Densidad de alojamiento
- Fotoperíodo

3. Variables nutricionales:

- Nivel de proteína en la dieta
- Suplementación mineral
- Sistema de alimentación

Ejemplo práctico detallado : En un experimento factorial 2×3 en cerdos de engorde, se evaluaron dos variables independientes:

- Factor A: Nivel de lisina digestible (0.9% vs 1.1%)
- Factor B: Densidad energética (3200, 3400, 3600 kcal EM/kg)

Se asignaron aleatoriamente 180 cerdos a 6 tratamientos (30 por tratamiento), evaluando el efecto sobre la ganancia diaria de peso (variable dependiente). El análisis mostró:

- Efecto principal lisina: +85 g/día con 1.1% ($p < 0.01$)
- Efecto principal energía: Respuesta cuadrática, óptimo en 3400 kcal
- Interacción lisina×energía: Significativa ($p = 0.03$), indicando que el efecto de la lisina dependía del nivel energético

b) Variables dependientes

La variable dependiente es la respuesta o resultado que se mide para evaluar el efecto de las variables independientes. Su selección debe basarse en relevancia biológica, mensurabilidad y sensibilidad al cambio.

Criterios de selección:

- Relevancia clínica o productiva
- Facilidad y precisión de medición
- Variabilidad conocida
- Correlación con outcomes de interés

Categorías principales:

1. **Respuestas clínicas:**

- Supervivencia
- Tiempo hasta recuperación
- Scores clínicos compuestos

2. **Respuestas productivas:**

- Ganancia de peso
- Producción láctea
- Conversión alimenticia

3. **Respuestas fisiológicas:**

- Parámetros hematológicos
- Niveles hormonales
- Marcadores de estrés

c) *Variables confusoras*

Las variables confusoras son factores que pueden influir tanto en la variable independiente como en la dependiente, potencialmente distorsionando la relación observada. Su identificación y control son cruciales para la validez interna del estudio.

Estrategias de control:

1. **En el diseño:**

- Aleatorización
- Restricción

- Apareamiento

2. En el análisis:

- Estratificación
- Ajuste multivariable
- Análisis de sensibilidad

Ejemplo práctico detallado: Un estudio evaluó la asociación entre obesidad y diabetes mellitus en gatos. La edad actuó como variable confusora, ya que:

- Gatos mayores tienen mayor prevalencia de obesidad ($r=0.42$)
- La diabetes es más frecuente en gatos geriátricos (OR=3.7 por cada 5 años)

Sin ajustar por edad, el OR crudo obesidad-diabetes fue 4.2 (IC 95%: 2.8-6.3). Después del ajuste multivariable por edad, el OR ajustado fue 2.9 (IC 95%: 1.9-4.4), demostrando que parte de la asociación inicial se debía al efecto confusor de la edad (Awe et al., 2022).

1.4 Poblaciones y muestras en estudios animales

1.4.1 Conceptos fundamentales de población

a) Definición

En investigación veterinaria, una población representa el conjunto completo de unidades (animales, granjas, muestras biológicas) que comparten características específicas y sobre las cuales se desea realizar inferencias. La definición precisa de la población objetivo constituye el

primer paso crítico en cualquier investigación, determinando la validez externa y aplicabilidad de los resultados.

b) Tipos de poblaciones en veterinaria

1. **Población Diana o Conceptual:** Representa el grupo ideal al cual se desean extrapolar los resultados. Por ejemplo, "todos los caninos con diabetes mellitus en América Latina". Esta población es teórica y raramente accesible en su totalidad.
2. **Población Accesible:** Subconjunto de la población diana que está disponible para el estudio. Continuando el ejemplo anterior, podría ser "caninos diabéticos atendidos en clínicas veterinarias de ciudades principales durante 2024".
3. **Población de Estudio:** Conjunto de individuos que cumplen los criterios de inclusión y exclusión definidos por el protocolo de investigación.

Ejemplo Práctico Detallado: En un estudio sobre prevalencia de *Leptospira* en bovinos lecheros:

- Población diana: Todos los bovinos lecheros de Ecuador
- Población accesible: Bovinos de 150 fincas registradas en asociaciones ganaderas de 5 provincias
- Población de estudio: 2,500 bovinos >2 años, de 75 fincas seleccionadas aleatoriamente
- Muestra final: 450 animales (6 por finca) seleccionados sistemáticamente

Esta jerarquización permite comprender las limitaciones en la generalización de resultados y la necesidad de describir claramente cada nivel poblacional en los reportes de investigación (King et al., 2024).

1.4.2 Características poblacionales en contexto veterinario

Las poblaciones animales presentan características únicas que influyen en el diseño de estudios:

1. Estructura jerárquica: Los animales frecuentemente están organizados en niveles jerárquicos naturales:

- Individual → Camada/Grupo → Corral/Jaula → Granja → Región → País
- Esta estructura requiere consideraciones especiales en muestreo y análisis

2. Dinamismo temporal: Las poblaciones animales son altamente dinámicas debido a:

- Nacimientos y muertes continuas
- Movimientos entre unidades productivas
- Cambios estacionales en tamaño y composición
- Ciclos productivos (entrada y salida de lotes)

3. Heterogeneidad intrínseca:

- Variabilidad genética entre razas y líneas
- Diferentes sistemas de manejo y producción
- Gradientes geográficos y climáticos

- Estados fisiológicos variables (gestación, lactancia, crecimiento)

Tabla 5. Consideraciones especiales según tipo de población animal.

Tipo de Población	Características Únicas	Desafíos de Muestreo	Estrategias de Recomendadas
Animales de Compañía	Población dispersa, vínculos emocionales con propietarios	Acceso limitado, sesgo de participación	Muestreo por conveniencia en clínicas, registros veterinarios
Animales de Producción	Agrupados en unidades, valor económico	Correlación intra-grupo, efectos de manejo	Muestreo por conglomerados, estratificación por sistema
Fauna Silvestre	Población móvil, difícil acceso	Detectabilidad variable, aspectos éticos	Captura-recaptura, muestreo adaptativo
Animales de Laboratorio	Población controlada, homogénea	Poca variabilidad, validez externa limitada	Aleatorización estricta, réplicas biológicas
Animales Zoológicos	Poblaciones pequeñas, alto valor individual	Tamaño muestral limitado, consideraciones de bienestar	Censo completo, estudios longitudinales

1.4.3 Conceptos y teoría del muestreo

a) Fundamentos del muestreo

El muestreo constituye el proceso de selección de un subconjunto representativo de la población que permita realizar inferencias válidas sobre el total. En medicina veterinaria, el muestreo adecuado es crucial tanto por razones prácticas (costos, logística) como éticas (minimizar el número de animales utilizados).

b) Propiedades de una buena muestra:

1. **Representatividad:** La muestra debe reflejar las características de la población en proporciones similares
2. **Adecuación:** Tamaño suficiente para detectar efectos de interés con poder estadístico apropiado
3. **Viabilidad:** Factible de obtener con los recursos disponibles
4. **Precisión:** Minimizar el error de muestreo mediante técnicas apropiadas

c) Tipos de muestreo en investigación veterinaria

1. Muestreo probabilístico

Muestreo aleatorio simple

Cada unidad de la población tiene igual probabilidad de ser seleccionada. Requiere un marco muestral completo (listado de todas las unidades).

Aplicación Veterinaria:

Ejemplo: Selección de 50 perros de un refugio con 500 animales

1. Asignar números del 1 al 500 a cada animal
2. Generar 50 números aleatorios usando software estadístico
3. Seleccionar los animales correspondientes

Ventajas:

- Elimina sesgo de selección
- Permite cálculo directo de error estándar
- Base para inferencia estadística

Limitaciones:

- Requiere marco muestral completo
- Puede ser logísticamente complejo
- Posible desbalance en características importantes

Muestreo sistemático

Selección de unidades a intervalos regulares después de un inicio aleatorio.

Ejemplo Práctico Detallado: En un matadero que procesa 1,200 bovinos/día, se requiere examinar 60 canales para detección de residuos.

Procedimiento:

1. Intervalo de muestreo $(k) = 1200/60 = 20$
2. Número aleatorio inicial entre 1 y 20: 7
3. Animales seleccionados: 7°, 27°, 47°, 67°... 1,187°

Este método es eficiente cuando los animales llegan en orden aleatorio respecto a la característica de interés. Sin embargo, debe evitarse si

existe periodicidad en el proceso (por ejemplo, si cada 20° animal proviene de la misma granja).

Muestreo estratificado

La población se divide en estratos homogéneos y se muestrea dentro de cada estrato.

Fórmulas para Asignación:

Asignación Proporcional: $n_{(i)} = n \times (N_i/N)$

Asignación Óptima de Neyman: $n_{(i)} = n \times (N_i \times \sigma_i) / \Sigma(N_i \times \sigma_i)$

Donde:

- $n_{(i)}$ = tamaño de muestra del estrato i
- n = tamaño de muestra total
- N_i = tamaño del estrato i
- σ_i = desviación estándar del estrato i

Ejemplo práctico detallado:

Tabla 6. Estudio de mastitis en 10,000 vacas lecheras estratificadas por número de lactancia.

Estrato (Lactancia)	N	Prevalencia Esperada	σ	Asignación Proporcional	Asignación Óptima
Primera	3,500	15%	0.357	105	95
Segunda	2,500	25%	0.433	75	82

Tercera	2,000	35%	0.477	60	72
Cuarta+	2,000	45%	0.497	60	75
Total	10,000	-	-	300	324

La asignación óptima asigna más muestras a estratos con mayor variabilidad, mejorando la precisión global de las estimaciones (Lima et al., 2024).

Muestreo por conglomerados

Se seleccionan grupos (conglomerados) completos en lugar de individuos.

Diseño de Dos Etapas:

1. Selección aleatoria de conglomerados (granjas)
2. Muestreo dentro de conglomerados seleccionados

Efecto del Diseño (DEFF): $DEFF = 1 + (m - 1) \times \rho$

Donde:

- m = tamaño promedio del conglomerado
- ρ = coeficiente de correlación intraclase

Ejemplo Práctico Detallado: Estudio de prevalencia de parásitos gastrointestinales en ovinos:

- 200 rebaños en la región (conglomerados)
- Selección de 20 rebaños aleatoriamente

- Muestreo de 15 animales por rebaño
- ICC estimado = 0.15 (alta similitud dentro de rebaños)
- $DEFF = 1 + (15-1) \times 0.15 = 3.1$

Esto significa que se requiere 3.1 veces más animales que en muestreo aleatorio simple para obtener la misma precisión. El tamaño de muestra efectivo es $300/3.1 = 97$ unidades independientes.

2. Muestreo no probabilístico

Muestreo por conveniencia

Selección basada en disponibilidad y accesibilidad.

Aplicaciones Válidas:

- Estudios piloto
- Validación de técnicas
- Generación de hipótesis
- Situaciones de emergencia sanitaria

Limitaciones:

- No permite inferencias poblacionales
- Alto riesgo de sesgo
- Validez externa limitada

Muestreo por cuotas

Similar al estratificado, pero sin selección aleatoria dentro de estratos.

Ejemplo veterinario: Estudio de comportamiento en 100 gatos domésticos:

- Cuota 1: 25 gatos <1 año
- Cuota 2: 50 gatos 1-7 años
- Cuota 3: 25 gatos >7 años Se incluyen los primeros casos disponibles hasta completar cada cuota.

Muestreo bolo de nieve

Útil para poblaciones de difícil acceso. Los participantes iniciales refieren a otros participantes potenciales.

Aplicación: Estudio de prácticas de manejo en criadores de especies exóticas. Los criadores registrados refieren a otros criadores no registrados, permitiendo acceder a población oculta.

1.4.4 Cálculo del tamaño de muestra

a) Factores Determinantes

El tamaño de muestra adecuado depende de múltiples factores que deben balancearse:

Parámetros estadísticos:

- Nivel de confianza ($1-\alpha$): Típicamente 95% ($\alpha = 0.05$)
- Poder estadístico ($1-\beta$): Mínimo 80%, idealmente 90%
- Tamaño del efecto: Diferencia mínima clínicamente relevante
- Variabilidad esperada: Basada en literatura o estudios piloto

Consideraciones Prácticas:

- Recursos disponibles (financieros, humanos, temporales)
- Aspectos éticos (principio de las 3R)
- Pérdidas esperadas durante el estudio
- Diseño del estudio (pareado, factorial, medidas repetidas)

b) Fórmulas según Tipo de Variable y Diseño

Para estimar una proporción:

$$n = Z^2(1-\alpha/2) \times p \times (1-p) / d^2$$

Donde:

- $Z(1-\alpha/2)$ = valor crítico de distribución normal (1.96 para 95% confianza)
- p = proporción esperada
- d = precisión deseada

Ejemplo cálculo: Prevalencia esperada de brucelosis bovina = 5%,
precisión = $\pm 2\%$ $n = 1.96^2 \times 0.05 \times 0.95 / 0.02^2 = 456$ animales

Para Comparar dos medias:

$$n = 2\sigma^2 \times (Z(1-\alpha/2) + Z(1-\beta))^2 / \delta^2$$

Donde:

- σ = desviación estándar común
- δ = diferencia mínima a detectar

Tabla 7. Tamaños de muestra para diferentes escenarios veterinarios.

Tipo de Estudio	Parámetro	Precisión/Efecto	n por grupo	Consideraciones
Prevalencia enfermedad rara	$p=1\%$	$\pm 0.5\%$	1,521	Requiere muestreo dirigido
Prevalencia común	$p=30\%$	$\pm 5\%$	323	Factible en la práctica
Ensayo clínico (continua)	$\mu_1 - \mu_2 = 0.5\sigma$	Poder 80%	64	Tamaño del efecto moderado
Ensayo clínico (proporción)	$p_1=70\%$, $p_2=90\%$	Poder 90%	87	Diferencia clínicamente importante
Estudio pareado	$d=0.3\sigma$	Poder 80%	90 pares	Reduce variabilidad inter-sujeto
Concordancia diagnóstica	κ esperado=0.7	Precisión ± 0.1	195	Para pruebas diagnósticas

Ajustes al Tamaño de Muestra

- Corrección por Población Finita:** Cuando $n > 5\%$ de N :

$$n(\text{ajustado}) = n / (1 + n/N)$$

2. **Ajuste por Pérdidas:** $n(\text{final}) = n / (1 - R)$

Donde R = tasa esperada de pérdidas

3. **Ajuste por Diseño Complejo:** $n(\text{complejo}) = n(\text{simple}) \times \text{DEFF}$

Ejemplo Integrado 1.13: Estudio de eficacia vacunal en pollos de engorde:

- Cálculo base: $n = 100$ por grupo
- Mortalidad esperada: 5% $\rightarrow n = 100/0.95 = 106$
- Diseño por conglomerados (galpones), $\text{DEFF} = 2 \rightarrow n = 106 \times 2 = 212$
- Total requerido: $212 \times 2 \text{ grupos} = 424$ pollos

1.4.5 Aspectos éticos del muestreo en animales

a) Principio de las 3R

Reemplazo (Replacement):

- Uso de modelos in vitro cuando sea posible
- Simulaciones computacionales para estudios preliminares
- Bases de datos históricos para análisis secundarios

Reducción (Reduction):

- Cálculo preciso del tamaño de muestra
- Diseños experimentales eficientes (crossover, bloques)

- Análisis interino para detener estudios tempranamente

Refinamiento (Refinement):

- Minimizar dolor y estrés durante muestreo
- Técnicas menos invasivas
- Enriquecimiento ambiental durante estudios

Ejemplo de Aplicación: En lugar de realizar eutanasia para obtener muestras de tejido hepático en 50 ratones, se implementó:

- Biopsia hepática guiada por ultrasonido (refinamiento)
- Reducción a 30 animales mediante diseño más eficiente
- Complementación con cultivos celulares para objetivos secundarios (reemplazo parcial)

1.5 Escalas de medición

1.5.1 Fundamentos teóricos de las escalas de medición

Las escalas de medición constituyen el sistema de reglas que define cómo se asignan números o categorías a las observaciones. En investigación veterinaria, la correcta identificación de la escala de medición determina no solo los análisis estadísticos aplicables, sino también la interpretación válida de los resultados. Stevens (1946) estableció la clasificación clásica en cuatro niveles que, aunque ha sido refinada, continúa siendo fundamental para la práctica estadística veterinaria.

1.5.2 Propiedades matemáticas de las escalas

Cada escala de medición posee propiedades matemáticas específicas que determinan las operaciones válidas:

- **Identidad:** Los números sirven como etiquetas únicas
- **Magnitud:** Los números representan cantidades ordenadas
- **Intervalos Iguales:** Las diferencias entre números son significativas
- **Cero Absoluto:** Existe un punto cero verdadero que indica ausencia total

1.5.3 Escala nominal en medicina veterinaria

a) Características y Propiedades

La escala nominal representa el nivel más básico de medición, donde los números o símbolos sirven únicamente como identificadores sin implicar orden o magnitud. En veterinaria, esta escala es fundamental para clasificar características cualitativas discretas.

b) Propiedades Matemáticas

- Solo posee la propiedad de identidad
- Los números son arbitrarios y sustituibles
- No admite operaciones aritméticas
- Solo permite determinar igualdad o diferencia

Ejemplo práctico detallado: En un estudio de asociación genómica (GWAS) en bovinos Holstein, se codificaron los genotipos del gen DGAT1 (asociado con contenido de grasa láctea):

- AA = 1 (homocigoto alelo A)
- AK = 2 (heterocigoto)
- KK = 3 (homocigoto alelo K)

Aunque se usan números, estos son meramente etiquetas. No significa que KK=3 sea "tres veces" AA=1. El análisis reveló que el genotipo KK incrementaba el contenido de grasa láctea en 0.23% comparado con AA ($p < 0.001$), utilizando pruebas apropiadas para datos nominales (Chi-cuadrado) (Ambarcıoğlu, 2021b).

1.5.4 Análisis estadístico de datos nominales

Medidas descriptivas válidas:

- Moda: Categoría más frecuente
- Frecuencias absolutas y relativas
- Razones y proporciones

Pruebas estadísticas aplicables:

- Chi-cuadrado de independencia
- Prueba exacta de Fisher (muestras pequeñas)
- Prueba de McNemar (datos pareados)
- V de Cramér (medida de asociación)

Tabla 8. Aplicaciones de escalas nominales en diferentes áreas veterinarias.

Área	Variable	Categorías	Análisis Típico	Interpretación Clínica
Genética	Grupo sanguíneo felino	A, B, AB	Frecuencias alélicas	Riesgo isoeritrolisis neonatal
Microbiología	Resistencia antibiótica	S, I, R	Proporción resistencia	Selección de tratamiento
Patología	Tipo histológico tumor	Carcinoma, Sarcoma, Linfoma	Distribución por tipo	Pronóstico diferencial
Nutrición	Fuente proteica dieta	Animal, Vegetal, Mixta	Asociación con respuesta	Formulación de raciones
Reproducción	Método inseminación	Natural, IA, TE	Comparación tasas éxito	Eficiencia reproductiva

1.5.5 Escala ordinal en medicina veterinaria

Características y aplicaciones

La escala ordinal añade la propiedad de orden a la escala nominal, permitiendo establecer relaciones de "mayor que" o "menor que" entre categorías, aunque sin cuantificar la magnitud de las diferencias.

Propiedades matemáticas:

- Posee identidad y magnitud
- El orden es significativo
- Las distancias entre categorías no son necesariamente iguales
- No permite operaciones aritméticas estándar

a) Escalas Ordinales Validadas en Veterinaria

1. Escalas de Dolor:

Glasgow Composite Pain Scale (Caninos):

- Componentes: Postura, vocalización, atención a herida, movilidad
- Rango: 0-24 puntos
- Punto de corte analgesia: >6 puntos
- Validación: Concordancia inter-observador $\kappa=0.85$

Ejemplo práctico detallado: En un estudio de 200 perros post-quirúrgicos, se evaluó el dolor cada 2 horas usando la escala Glasgow. Los resultados mostraron:

- Mediana pre-analgesia: 9 (RIQ: 7-12)
- Mediana post-analgesia: 3 (RIQ: 2-5)
- Prueba de Wilcoxon: $Z = -8.34$, $p < 0.001$

El uso de medianas y rangos intercuartílicos es apropiado para datos ordinales, evitando la tentación incorrecta de calcular medias que asumirían intervalos iguales entre puntos de la escala (Ambarcıoğlu, 2021).

2. Sistemas de Clasificación Clínica:

Tabla 9. Sistema IRIS para Enfermedad Renal Crónica.

Estadio	Creatinina (mg/dL)		Descripción	Manejo Recomendado
1	<1.4 (gato) (perro)	<1.4	No azotémico	Monitoreo, identificar causa
2	1.4-2.8 (gato) (perro)	1.4-2.0	Azotemia leve	Dieta renal, control presión
3	2.9-5.0 (gato) (perro)	2.1-5.0	Azotemia moderada	Terapia intensiva, fluidoterapia
4	>5.0		Azotemia severa	Cuidados paliativos, diálisis

Aunque basado en valores continuos de creatinina, el sistema IRIS crea categorías ordinales que guían decisiones terapéuticas escalonadas.

b) Análisis Estadístico de Datos Ordinales

Medidas de tendencia central:

- Mediana: Valor central cuando datos se ordenan
- Moda: Categoría más frecuente
- NO usar media aritmética (asume intervalos iguales)

Medidas de dispersión:

- Rango intercuartílico (RIQ)
- Rango

- NO usar desviación estándar

Pruebas estadísticas:

- Mann-Whitney U (comparar 2 grupos independientes)
- Wilcoxon (datos pareados)
- Kruskal-Wallis (>2 grupos)
- Correlación de Spearman
- Regresión ordinal

1.5.6 Escala de intervalo en medicina veterinaria

a) Características fundamentales

La escala de intervalo posee intervalos iguales entre puntos consecutivos, pero carece de un cero absoluto verdadero. El cero es arbitrario y no indica ausencia total del atributo medido.

b) Propiedades matemáticas

- Identidad, magnitud e intervalos iguales
- Permite sumas y restas
- No permite multiplicación/división significativa
- Las razones no son interpretables

c) Aplicaciones veterinarias

Temperatura corporal: El ejemplo clásico en medicina veterinaria es la temperatura en grados Celsius o Fahrenheit.

Ejemplo práctico detallado: Estudio de variación circadiana de temperatura en equinos:

- Temperatura matutina promedio: 37.2°C
- Temperatura vespertina promedio: 38.1°C
- Diferencia: 0.9°C

Esta diferencia de 0.9°C es significativa y constante. Sin embargo, no podemos decir que 38.1°C es "1.024 veces más caliente" que 37.2°C, pues el cero en Celsius es arbitrario. La misma diferencia en Kelvin sería 311.25K vs 310.35K, dando una razón diferente (1.003).

Scores Compuestos: Muchas escalas clínicas compuestas se tratan como intervalos, aunque sean técnicamente ordinales.

Tabla 10. Escalas de intervalo y cuasi-intervalo en veterinaria.

Escala	Rango	Aplicación	Consideraciones
Temperatura (°C)	Variable	Fisiología, termorregulación	Verdadera escala de intervalo
pH sanguíneo	6.8-7.8	Equilibrio ácido-base	Escala logarítmica, tratada como intervalo
Score APACHE veterinario	0-100	Pronóstico UCI	Suma ponderada, cuasi-intervalo
Índice de masa corporal	Variable	Evaluación nutricional	Derivado de medidas de razón

1.5.7 Escala de razón en medicina veterinaria

a) Características y propiedades

La escala de razón representa el nivel más completo de medición, con todas las propiedades de las escalas anteriores más un cero absoluto verdadero que indica ausencia total del atributo.

b) Propiedades matemáticas:

- Todas las propiedades: identidad, magnitud, intervalos iguales, cero absolutos
- Permite todas las operaciones aritméticas
- Las razones son significativas
- Transformaciones permisibles: multiplicación por constante

c) Variables de razón en veterinaria

1. Medidas Físicas:

- Peso corporal (kg)
- Longitud (cm)
- Volumen (mL)
- Tiempo de supervivencia (días)

2. Concentraciones:

- Glucemia (mg/dL)
- Recuento celular (células/ μ L)
- Concentración hormonal (ng/mL)
- Carga parasitaria (HPG)

3. Medidas Productivas:

- Producción láctea (L/día)
- Ganancia de peso (g/día)

- Conversión alimenticia (kg alimento/kg ganancia)

Ejemplo práctico detallado:

Tabla 11. Evaluación de crecimiento en corderos con diferentes dietas.

Tratamiento	Peso Inicial (kg)	Peso Final (kg)	GDP (g/día)	Conversión
Control	22.5 ± 2.1	38.7 ± 3.2	180 ± 21	4.8:1
Suplementado	22.8 ± 1.9	45.3 ± 2.8	250 ± 18	3.5:1

Las razones son interpretables:

- Los corderos suplementados ganaron $250/180 = 1.39$ veces más peso/día
- La conversión mejoró en 27% (3.5 vs 4.8)
- El peso final del grupo suplementado fue 17% mayor

Todas estas comparaciones son válidas porque las variables tienen cero absoluto real.

d) Transformaciones y Análisis

Transformaciones Válidas:

- Logarítmica (normalizar distribuciones sesgadas)
- Raíz cuadrada (estabilizar varianza)
- Recíproca (cambiar dirección de relación)
- Box-Cox (optimizar normalidad)

Ejemplo de Transformación: Concentración de cortisol con distribución log-normal:

- Datos originales: 15, 22, 45, 120, 280 ng/mL
- Media geométrica: $\exp(\text{mean}(\log(\text{datos}))) = 52.4 \text{ ng/mL}$
- Más representativa que media aritmética: 96.4 ng/mL

1.5.8 Consideraciones prácticas y errores comunes

a) Selección apropiada de escala

Principios guía:

1. Usar la escala más informativa posible
2. No asumir propiedades que no existen
3. Considerar el objetivo del análisis
4. Validar supuestos antes del análisis

Errores frecuentes en veterinaria:

1. **Tratar Datos Ordinales como Intervalo:**
 - Error: Calcular media de scores de condición corporal
 - Correcto: Usar mediana y pruebas no paramétricas
2. **Dicotomización Innecesaria:**
 - Error: Convertir glucemia en "normal/alto"
 - Pérdida de información y poder estadístico
 - Mejor: Mantener variable continua
3. **Análisis Inapropiado de Porcentajes:**
 - Error: Aplicar prueba t a porcentajes cerca de 0 o 100%

- Correcto: Transformación arcoseno o modelos binomiales

Tabla 12. Guía de selección de análisis según escala de medición.

Escala	Tendencia Central	Dispersión	Correlación	Prueba 2 Grupos	Prueba >2 Grupos
Nominal	Moda	Índice variación	Phi, V de Cramér	Chi-cuadrado, Fisher	Chi-cuadrado
Ordinal	Mediana	RIQ, Rango	Spearman, Kendall	Mann-Whitney U	Kruskal-Wallis
Intervalo	Media	DE, Varianza	Pearson	t de Student	ANOVA
Razón	Media, Media geométrica	DE, CV	Pearson	t de Student	ANOVA

1.6 Organización y presentación de datos

1.6.1 Principios de organización de datos

a) Estructura de bases de datos

La organización sistemática de datos constituye el fundamento para cualquier análisis estadístico válido. En investigación veterinaria, donde

los datos provienen de fuentes diversas (registros clínicos, laboratorio, producción), establecer una estructura coherente desde el inicio del estudio es crucial.

b) Principios fundamentales

1. Formato Rectangular (Tidy Data):

- Cada variable forma una columna
- Cada observación forma una fila
- Cada tipo de unidad observacional forma una tabla

```
| ID_Animal | Fecha | Peso_kg | Temp_C | Tratamiento | Granja |
|-----|-----|-----|-----|-----|-----|
| B001 | 2024-01-15 | 45.2 | 38.5 | Control | Granja_A |
| B002 | 2024-01-15 | 48.7 | 38.8 | Vacunado | Granja_A |
| B003 | 2024-01-15 | 44.1 | 38.6 | Control | Granja_B |
```

Figura 1. Ejemplo de estructura correcta.

2. Codificación Consistente:

- Usar códigos estandarizados para categorías
- Mantener diccionario de datos actualizado
- Evitar espacios y caracteres especiales en nombres de variables

3. Manejo de Datos Faltantes:

- Distinguir entre tipos de datos faltantes (MCAR, MAR, MNAR)
- Usar códigos específicos (NA, -999) consistentemente
- Documentar razones de datos faltantes

Ejemplo práctico detallado: En un estudio longitudinal de crecimiento en pollos de engorde con 1,000 aves seguidas durante 42 días:

Estructura de datos:

- Nivel 1: Mediciones (peso, consumo) - 7,000 registros
- Nivel 2: Animal (ID, sexo, línea genética) - 1,000 registros
- Nivel 3: Corral (tratamiento, temperatura) - 40 registros
- Nivel 4: Galpón (ventilación, orientación) - 4 registros

Esta estructura jerárquica requirió tablas relacionadas mediante códigos únicos, permitiendo análisis multinivel que consideró la variabilidad en cada nivel (Paredes-Catota et al., 2021).

c) Control de calidad de datos

Procedimientos de validación:

1. Verificación de rangos:

```
Ejemplo en R:  
# Verificar valores fisiológicos válidos  
datos %>%  
filter(temperatura < 36 | temperatura > 42) %>%  
flag_anomaly()
```

Figura 2. Verificación de rangos.

2. Consistencia lógica:

- Fecha de muerte no puede preceder fecha de nacimiento
- Peso al destete debe ser mayor que peso al nacimiento
- Producción láctea requiere estado "lactante"

3. Detección de valores atípicos:

- Método del rango intercuartílico (RIQ)
- Z-scores para variables normales
- Distancia de Mahalanobis para datos multivariados

Tabla 13. Protocolo de validación de datos veterinarios.

Tipo de Error	Método de Detección	de Acción Correctiva	Ejemplo
Fuera de rango	Límites biológicos	Verificar fuente, corregir o eliminar	Frecuencia cardíaca >300 en bovino
Inconsistencia temporal	Ordenamiento cronológico	Revisar fechas	Destete antes del nacimiento
Duplicados	Identificadores únicos	Eliminar o consolidar	Mismo ID en diferentes fechas simultáneas
Datos imposibles	Reglas lógicas	Investigar y corregir	Macho gestante
Outliers extremos	$>3 \text{ DE}$ o $>3 \times \text{RIQ}$	Verificar, transformar o análisis robusto	Peso 500 kg en ovino

1.6.2 Tablas de frecuencia y contingencia

a) Tablas de frecuencia univariadas

Las tablas de frecuencia constituyen la forma más básica de resumir datos categóricos, proporcionando información sobre la distribución de una variable.

Componentes esenciales:

- Frecuencia absoluta (n)
- Frecuencia relativa (%)
- Frecuencia acumulada
- Intervalos de confianza para proporciones

Ejemplo práctico detallado:

Tabla 14. Distribución de diagnósticos en 500 consultas de emergencia veterinaria.

Categoría Diagnóstica	n	%	% Acumulado	IC 95%
Trauma	142	28.4	28.4	24.5-32.6
Gastrointestinal	118	23.6	52.0	20.0-27.6
Respiratorio	75	15.0	67.0	12.0-18.5
Intoxicación	68	13.6	80.6	10.8-17.0
Neurológico	52	10.4	91.0	7.9-13.5
Otros	45	9.0	100.0	6.7-11.9

Total	500	100.0	-	-
--------------	------------	--------------	----------	----------

Los intervalos de confianza se calcularon usando el método Wilson para proporciones binomiales, más preciso que el método normal para proporciones extremas.

b) Tablas de Contingencia

Las tablas de contingencia (tablas cruzadas) muestran la distribución conjunta de dos o más variables categóricas.

Estructura y análisis:

Ejemplo práctico detallado:

Tabla 15. Relación entre método diagnóstico y resultado en detección de Brucella en 300 bovinos.

Método	Positivo	Negativo	Total	Sensibilidad	Especificidad
Rosa de Bengala	45	255	300	90.0%	94.1%
ELISA	48	252	300	96.0%	93.3%
PCR	50	250	300	100.0%	92.6%
Real (Gold Standard)	50	250	300	-	-

Concordancia Rosa Bengala vs ELISA:			
	ELISA+	ELISA-	
Rosa Bengala+	42	3	45
Rosa Bengala-	6	249	255
	48	252	300
Kappa = 0.85 (IC 95%: 0.77-0.93)			
Interpretación: Concordancia casi perfecta			

Figura 3. Análisis de Concordancia.

Distribuciones de frecuencia para variables continuas

Proceso de agrupación:

- Determinar número de clases (k):**
 - Regla de Sturges: $k = 1 + 3.322 \times \log_{10}(n)$
 - Regla de la raíz cuadrada: $k = \sqrt{n}$
 - Regla de Rice: $k = 2 \times n^{(1/3)}$
- Calcular amplitud de clase:**
 - Amplitud = (Valor máximo - Valor mínimo) / k

Ejemplo práctico detallado:

Tabla 16. Distribución de pesos al sacrificio de 200 cerdos.

Clase (kg)	Peso Marca Clase	Frecuencia	%	% Acumulado	Densidad
90-95	92.5	8	4.0	4.0	0.008
95-100	97.5	22	11.0	15.0	0.022
100-105	102.5	48	24.0	39.0	0.048

105-110	107.5	62	31.0	70.0	0.062
110-115	112.5	38	19.0	89.0	0.038
115-120	117.5	16	8.0	97.0	0.016
120-125	122.5	6	3.0	100.0	0.006

Nota. Media = 107.3 kg, Mediana = 107.1 kg, Moda = 107.5 kg La distribución es aproximadamente normal, validada por prueba de Shapiro-Wilk ($p = 0.23$).

1.6.3 Representación gráfica de datos

a) Principios de visualización efectiva

Criterios de excelencia gráfica (Tufte, 2001):

- Mostrar los datos
- Inducir a pensar sobre el contenido
- Evitar distorsiones
- Presentar muchos números en poco espacio
- Hacer coherentes grandes conjuntos de datos
- Revelar datos a varios niveles de detalle

b) Gráficos para variables categóricas

1. Gráficos de barras:

Apropiados para frecuencias de variables nominales u ordinales.

```
# Gráfico de barras con intervalos de confianza
library(ggplot2)
ggplot(data_diagnosticos, aes(x = categoria, y = porcentaje)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  geom_errorbar(aes(ymin = lower_ci, ymax = upper_ci), width = 0.2) +
  labs(title = "Distribución de Diagnósticos de Emergencia",
        x = "Categoría Diagnóstica",
        y = "Porcentaje de Casos") +
  theme_minimal()
```

Figura 4. Ejemplo Práctico

2. Gráficos de sectores (Pie Charts):

Útiles solo cuando se muestra composición de un todo.

Limitaciones:

- Difícil comparar sectores similares
- No apropiados para muchas categorías
- Preferir gráficos de barras en publicaciones científicas

3. Gráficos de mosaico:

Excelentes para visualizar tablas de contingencia complejas.

c) Gráficos para variables continuas

1. Histogramas:

Muestran la distribución de frecuencias de una variable continua.

Consideraciones:

- Número de bins afecta interpretación
- Agregar curva de densidad para suavizar

- Comparar con distribución teórica

```
# Histograma con curva normal superpuesta
hist(peso_cerdos, breaks = 30, probability = TRUE,
     main = "Distribución de Pesos al Sacrificio",
     xlab = "Peso (kg)", col = "lightgray")
curve(dnorm(x, mean = mean(peso_cerdos), sd = sd(peso_cerdos)),
     add = TRUE, col = "red", lwd = 2)
```

Figura 5. Ejemplo Código R.

2. Diagramas de caja (Boxplots):

Resumen visual de cinco números: mínimo, Q1, mediana, Q3, máximo.

Ventajas:

- Identifican outliers automáticamente
- Permiten comparar múltiples grupos
- Muestran asimetría de distribución

Ejemplo práctico detallado:

Estado	Min	Q1	Med	Q3	Max	Outliers
Preparto	45	52	58	65	78	92, 95
Posparto	28	35	42	48	61	22, 25, 72
Lactancia	38	48	55	62	75	None

Figura 6. Comparación de glucemia en tres estados metabólicos de vacas lecheras.

El boxplot reveló hipoglucemia marcada en el período posparto inmediato, información crucial para el manejo nutricional.

3. Gráficos de dispersión:

Visualizan la relación entre dos variables continuas.

Elementos Importantes:

- Línea de tendencia (regresión)
- Intervalos de confianza/predicción
- Identificación de puntos influyentes

4. Gráficos de densidad:

Versión suavizada del histograma usando estimación kernel.

Tabla 17. Selección de gráficos según tipo de datos y objetivo.

Tipo de Datos		Objetivo	Gráfico Recomendado	Alternativas
Categorica única		Mostrar frecuencias	Barras	Lollipop plot
Categorica	vs	Asociación	Mosaico	Barras agrupadas
Continua única		Distribución	Histograma	Densidad, Q-Q plot
Continua	vs	Comparar grupos	Boxplot	Violin plot, Ridgeline
Continua	vs	Correlación	Dispersión	Hexbin (muchos datos)
Serie temporal		Tendencia	Líneas	Área, Suavizado

Multivariada	Patrones	PCA biplot	Heatmap, Parallel coordinates
--------------	----------	------------	-------------------------------

d) *Gráficos especializados en veterinaria*

1. Curvas de supervivencia (Kaplan-Meier):

Esenciales en oncología veterinaria y estudios de longevidad.

Ejemplo práctico detallado:

Tiempo (días)	CHOP(%)	COP(%)	HR	p-valor
30	95	88	-	-
60	82	70	0.68	0.042
90	68	51	0.62	0.018
180	45	28	0.58	0.009
365	22	10	0.55	0.003

Mediana supervivencia: CHOP = 162 días, COP = 94 días
Log-rank test: $\chi^2 = 8.74$, $p = 0.003$

Figura 7. Supervivencia de perros con linfoma según protocolo de quimioterapia.

2. Gráficos de Bland-Altman:

Evalúan concordancia entre dos métodos de medición.

3. Forest Plots:

Resumen resultados de meta-análisis o múltiples comparaciones.

1.6.4 Medidas de resumen y estadísticos descriptivos

a) Medidas de tendencia central

1. Media Aritmética:

$$\bar{x} = \Sigma x_i / n$$

Propiedades:

- Sensible a valores extremos
- Óptima para distribuciones simétricas
- Permite inferencias paramétricas

2. Mediana:

Valor que divide la distribución en dos mitades iguales.

Ventajas:

- Robusta a outliers
- Apropiaada para distribuciones asimétricas
- Interpretación intuitiva
- 3. Moda:

Valor más frecuente en la distribución.

Aplicaciones:

- Variables categóricas
- Identificar picos en distribuciones multimodales
- Tamaño de camada más común

Ejemplo práctico comparativo: Concentración de células somáticas en leche ($\times 1000$ células/mL):

- Media: 285 (influenciada por valores altos)
- Mediana: 180 (mejor representación central)
- Moda: 150 (valor más frecuente)
- Media geométrica: 195 (apropiada para datos log-normales)

b) *Medidas de Dispersión*

Rango: Diferencia entre valor máximo y mínimo.

Varianza: $s^2 = \Sigma(x_i - \bar{x})^2 / (n-1)$

Desviación Estándar: $s = \sqrt{s^2}$

Coefficiente de Variación: $CV = (s/\bar{x}) \times 100\%$

Interpretación en Veterinaria:

- CV <10%: Baja variabilidad
- CV 10-30%: Variabilidad moderada
- CV >30%: Alta variabilidad

Error Estándar: $SE = s/\sqrt{n}$

Tabla 18. Medidas de dispersión según escala y distribución.

Medida	Escala	Distribución	Uso Principal	Ejemplo
	Requerida			Veterinario

Rango	Ordinal+	Cualquiera	Descripción simple	Rango temperaturas diarias
RIQ	Ordinal+	Cualquiera	Robusta outliers	a Score condición corporal
Varianza	Intervalo+	Preferible normal	Base para inferencias	Peso vivo en población
DE	Intervalo+	Preferible normal	Interpretación directa	Glucemia en ayunas
CV	Razón	Cualquiera positiva	Comparar variabilidades	Producción láctea vs peso
MAD	Ordinal+	Cualquiera	Muy robusta	Conteos parasitarios

c) Medidas de Forma

1. Asimetría (Skewness):

$$g_1 = [\Sigma(x_i - \bar{x})^3/n]/s^3$$

Interpretación:

- $g_1 = 0$: Simétrica
- $g_1 > 0$: Asimetría positiva (cola derecha)
- $g_1 < 0$: Asimetría negativa (cola izquierda)

2. Curtosis:

$$g_2 = [\Sigma(x_i - \bar{x})^4/n]/s^4 - 3$$

Interpretación:

- $g_2 = 0$: Mesocúrtica (normal)
- $g_2 > 0$: Leptocúrtica (colas pesadas)
- $g_2 < 0$: Platicúrtica (colas ligeras)

1.6.5 Presentación profesional de resultados

a) Estándares para tablas científicas

Formato APA (7ª edición) para Tablas:

1. Número de tabla en negrita
2. Título descriptivo en cursiva
3. Encabezados de columna claros
4. Cuerpo de tabla con datos
5. Notas al pie si necesario

```

# Análisis completo de datos veterinarios
library(tidyverse)
library(epiR)

# Cargar y explorar datos
datos <- read.csv("mastitis_bovinos.csv")
summary(datos)

# Calcular prevalencia con IC
prev <- epi.prev(pos = 45, tested = 300,
                 conf.level = 0.95, method = "wilson")

# Visualización
ggplot(datos, aes(x = num_lactancia, y = scc_log)) +
  geom_boxplot(aes(fill = factor(num_lactancia))) +
  scale_y_log10() +
  theme_minimal() +
  labs(title = "Células Somáticas por Número de Lactancia",
       x = "Número de Lactancia",
       y = "SCC (células/mL, escala log)")

```

Figura 8. Ejemplo de Código R para análisis veterinario.

1.6.6 El método científico en investigación veterinaria

a) Etapas del proceso de investigación estadística

Formulación del problema: Identificar claramente la pregunta de investigación en términos medibles y específicos.

Ejemplo Desarrollado:

- Pregunta vaga: "¿Funciona el nuevo antiparasitario?"
- Pregunta específica: "¿El antiparasitario X reduce el conteo de huevos por gramo de heces (HPG) en ovinos naturalmente

infectados con *Haemonchus contortus* en al menos 90% comparado con el grupo control, 14 días post-tratamiento?"

Diseño del estudio: Seleccionar el diseño apropiado considerando:

- Objetivos (descriptivo, analítico, experimental)
- Recursos disponibles
- Consideraciones éticas
- Validez interna y externa

Recolección de datos: Implementar protocolos estandarizados asegurando:

- Calibración de instrumentos
- Entrenamiento de personal
- Cegamiento cuando sea apropiado
- Cadena de custodia de muestras

Análisis estadístico: Aplicar métodos apropiados según:

- Tipo de variables
- Distribución de datos
- Supuestos del modelo
- Objetivos del estudio

Interpretación y conclusiones: Traducir resultados estadísticos a conclusiones biológicas/clínicas considerando:

- Significancia estadística vs relevancia clínica
- Limitaciones del estudio
- Implicaciones prácticas
- Direcciones futuras

b) Ejercicios y problemas

Ejercicios resueltos

Ejercicio 1: Identificación de variables

En un estudio sobre eficacia de un antiparasitario en ovinos, identifique el tipo de variable y escala de medición para cada una:

- a) Peso corporal (kg) b) Grado de anemia (leve/moderado/severo) c) Conteo de huevos por gramo de heces d) Raza del animal e) Temperatura rectal (°C) f) Presencia de diarrea (sí/no)

Solución: a) Cuantitativa continua, escala de razón b) Cualitativa ordinal, escala ordinal c) Cuantitativa discreta, escala de razón d) Cualitativa nominal, escala nominal e) Cuantitativa continua, escala de intervalo f) Cualitativa nominal (dicotómica), escala nominal.

Ejercicio 2: Cálculo de tamaño de muestra

Se desea estimar la prevalencia de brucelosis bovina en una región. Estudios previos sugieren una prevalencia del 8%. ¿Cuántos animales deben muestrearse para obtener una precisión de $\pm 2\%$ con 95% de confianza?

Solución: $n = Z^2_{(1-\alpha/2)} \times p \times (1-p) / d^2$ $n = 1.96^2 \times 0.08 \times 0.92 / 0.02^2$ $n = 3.8416 \times 0.0736 / 0.0004$ $n = 707$ animales

Considerando 10% de pérdidas: $n = 707/0.9 = 786$ animales

Problemas propuestos

Nivel básico

1. Clasifique las siguientes variables según su tipo y escala de medición:
 - Número de partos de una vaca
 - Color del pelaje en caballos
 - Concentración de proteína en alimento balanceado
 - Grado de claudicación (0-5)
 - Tiempo de recuperación post-anestesia
2. En una clínica veterinaria se registraron 150 consultas. Construya una tabla de frecuencias si las categorías fueron: vacunación (45), consulta médica (62), cirugía (28), emergencia (15).

3. Calcule media, mediana y moda para los siguientes pesos de lechones al destete (kg): 5.8, 6.2, 5.9, 7.1, 6.8, 5.9, 6.5, 7.0, 6.2, 5.9

Nivel intermedio

4. Un estudio compara dos tratamientos para mastitis bovina. Si la tasa de curación del tratamiento estándar es 70% y se espera que el nuevo tratamiento alcance 85%, calcule el tamaño de muestra necesario por grupo ($\alpha=0.05$, $\beta=0.20$).
5. Se tienen datos de producción láctea de 100 vacas Holstein. Los datos muestran una media de 28 L/día con desviación estándar de 4.5 L/día. a) Calcule el coeficiente de variación b) Determine el intervalo que contiene el 95% de las observaciones (asumiendo normalidad) c) Si se seleccionan 10 vacas al azar, ¿cuál es el error estándar de la media?
6. Diseñe un plan de muestreo estratificado para evaluar la condición corporal en un rebaño de 500 cabras divididas en: cabritas (120), adultas no gestantes (200), gestantes (100), y lactantes (80). La muestra total debe ser de 60 animales.

Nivel avanzado

7. En un experimento factorial 2×2 se evalúa el efecto del tipo de alojamiento (individual vs. grupal) y la suplementación vitamínica (con vs. sin) sobre la ganancia de peso en cerdos. Diseñe la estructura de datos necesaria para 40 animales (10 por

tratamiento) con mediciones semanales durante 8 semanas. Incluya variables de identificación, tratamiento y respuesta.

8. Un veterinario sospecha que existe asociación entre el método de parto (natural vs. cesárea) y la supervivencia neonatal en cachorros. De 200 partos observados:

- Parto natural: 150 casos, 12 muertes neonatales
- Cesárea: 50 casos, 8 muertes neonatales

Construya la tabla de contingencia, calcule las tasas de mortalidad y determine si existe asociación significativa.

9. Desarrolle un protocolo completo de organización de datos para un estudio longitudinal de 6 meses sobre el efecto de tres dietas en el control de diabetes felina, considerando:

- 30 gatos diabéticos
- Mediciones mensuales de glucemia, fructosamina y peso
- Registro de eventos adversos
- Estructura de base de datos
- Plan de análisis estadístico

1.7 Conclusiones del capítulo

Este capítulo ha establecido los fundamentos esenciales de la bioestadística aplicada a las ciencias veterinarias, consolidando las bases conceptuales y metodológicas necesarias para un análisis riguroso de los datos en el contexto de la salud y la producción animal. Se ha evidenciado que la bioestadística es una herramienta indispensable en

la práctica veterinaria contemporánea, ya que sustenta la toma de decisiones basadas en evidencia tanto en el ámbito clínico como en el productivo.

La correcta identificación y clasificación de las variables, junto con el reconocimiento de sus respectivas escalas de medición (nominal, ordinal, de intervalo y de razón), resulta determinante para seleccionar los procedimientos estadísticos adecuados y garantizar la validez de las interpretaciones. Asimismo, se ha resaltado la importancia de comprender las diferencias entre poblaciones y muestras, destacando que un diseño de muestreo apropiado asegura la representatividad de los datos y la solidez de las inferencias, especialmente en poblaciones animales con características particulares.

De igual forma, la adecuada organización, sistematización y presentación de la información constituye un paso fundamental no solo para el análisis estadístico, sino también para la comunicación científica clara y efectiva. En conjunto, estos elementos conforman el marco teórico-práctico que permite al profesional veterinario abordar con rigor y precisión los desafíos investigativos y clínicos, promoviendo una práctica fundamentada en la evidencia cuantitativa.

1.8 Autoevaluación del capítulo

Preguntas de comprensión:

1. ¿Por qué es importante distinguir entre variables cuantitativas continuas y discretas en el análisis veterinario?

2. Explique con un ejemplo veterinario la diferencia entre población diana, población accesible y muestra.
3. ¿Cuándo es apropiado usar la mediana en lugar de la media como medida de tendencia central?
4. Describa tres situaciones en medicina veterinaria donde el muestreo estratificado sería preferible al muestreo aleatorio simple.
5. ¿Qué consideraciones éticas deben tenerse en cuenta al determinar el tamaño de muestra en estudios con animales?

CAPÍTULO II

2 ESTADÍSTICA DESCRIPTIVA

2.1 Objetivos de aprendizaje

Al finalizar este capítulo, el estudiante será capaz de:

1. Calcular e interpretar las principales medidas de tendencia central en datos veterinarios
2. Aplicar medidas de dispersión para evaluar la variabilidad en parámetros clínicos animales
3. Analizar la forma y posición de distribuciones de datos biológicos
4. Construir e interpretar representaciones gráficas apropiadas para diferentes tipos de variables veterinarias
5. Implementar análisis descriptivos usando R en situaciones clínicas reales
6. Seleccionar las medidas descriptivas más apropiadas según el tipo de variable y distribución de los datos

2.2 Introducción

La estadística descriptiva constituye el primer paso fundamental en el análisis de datos en ciencias veterinarias. Antes de realizar inferencias sobre poblaciones animales o comparar tratamientos experimentales, resulta esencial comprender y describir adecuadamente las características de los datos recolectados. En la práctica veterinaria, la estadística descriptiva permite resumir información compleja de manera

comprensible, facilitando la toma de decisiones clínicas basadas en evidencia.

Este capítulo presenta las herramientas estadísticas fundamentales para describir y visualizar datos veterinarios, desde parámetros hematológicos hasta medidas productivas en sistemas ganaderos. Se enfatiza la interpretación práctica de cada medida estadística y su aplicación directa en contextos veterinarios reales, proporcionando las bases para análisis más complejos en capítulos posteriores.

2.3 Medidas de tendencia central

2.3.1 Conceptos fundamentales

Las medidas de tendencia central representan valores típicos o representativos de un conjunto de datos, permitiendo caracterizar el centro de una distribución. En veterinaria, estas medidas resultan fundamentales para establecer valores de referencia, evaluar el estado de salud de poblaciones animales y comparar grupos bajo diferentes condiciones (Alvitez-Temoche et al., 2024).

2.3.2 Media aritmética

La media aritmética ($\bar{x}$) representa el promedio de un conjunto de observaciones y constituye la medida de tendencia central más utilizada en investigación veterinaria.

Fórmula:

$$\bar{x} = \sum x_i / n$$

Donde:

- x_i = cada observación individual
- n = número total de observaciones
- Σ = sumatoria

Interpretación Veterinaria: La media proporciona un valor representativo cuando los datos presentan distribución simétrica y no existen valores extremos. En medicina veterinaria, se utiliza frecuentemente para:

- Establecer valores hematológicos de referencia
- Determinar pesos promedio en lotes de producción
- Calcular tasas de crecimiento en animales de engorde

Ejemplo práctico: Glucemia en caninos

Un veterinario registró los niveles de glucosa sanguínea (mg/dL) en 12 perros aparentemente sanos:

Datos: 85, 92, 88, 95, 91, 89, 94, 87, 90, 93, 86, 91

Media = $(85 + 92 + 88 + 95 + 91 + 89 + 94 + 87 + 90 + 93 + 86 + 91) / 12 = 90.08$ mg/dL

Interpretación: El nivel promedio de glucosa en este grupo es 90.08 mg/dL, valor que se encuentra dentro del rango normal para caninos (70-110 mg/dL).

2.3.3 Mediana

La mediana (Me) representa el valor que divide al conjunto de datos ordenados en dos mitades iguales. Esta medida resulta especialmente útil cuando existen valores atípicos o distribuciones asimétricas.

Cálculo:

1. Ordenar los datos de menor a mayor
2. Si n es impar: $Me = \text{valor central}$
3. Si n es par: $Me = \text{promedio de los dos valores centrales}$

Aplicaciones veterinarias:

- Análisis de tiempos de recuperación postquirúrgica
- Evaluación de períodos de incubación de enfermedades
- Determinación de edades en poblaciones animales con individuos muy longevos

Ejemplo práctico: Tiempo de recuperación anestésica en felinos

Tiempos de recuperación (minutos) en 15 gatos sometidos a ovariectomía:

Datos ordenados: 12, 14, 15, 16, 18, 19, 20, 22, 23, 25, 28, 32, 35, 42, 65

Posición de la mediana = $(15 + 1) / 2 = 8$ Mediana = 22 minutos

Interpretación: El 50% de los gatos se recuperó en 22 minutos o menos. Note que la media (25.73 minutos) se ve afectada por el valor

extremo de 65 minutos, mientras que la mediana proporciona una mejor representación del tiempo típico de recuperación.

2.3.4 Moda

La moda (Mo) es el valor que aparece con mayor frecuencia en un conjunto de datos. Un conjunto puede ser unimodal, bimodal o multimodal según el número de modas presentes.

Relevancia Veterinaria:

- Identificación de diagnósticos más frecuentes en clínicas
- Determinación de parásitos prevalentes en una región
- Análisis de patrones reproductivos estacionales

Ejemplo práctico: Diagnósticos en urgencias veterinarias

Tabla 19. Registro mensual de casos de urgencia por categoría diagnóstica.

Diagnóstico	Frecuencia
Traumatismo	45
Intoxicación	32
Gastroenteritis	58
Distrés respiratorio	28
Convulsiones	15

Moda = Gastroenteritis (58 casos)

Interpretación: La gastroenteritis representa la causa más frecuente de consulta de urgencia, información crucial para la planificación de recursos y capacitación del personal.

2.3.5 Media ponderada

La media ponderada asigna diferentes pesos o importancias a cada observación, siendo particularmente útil cuando los datos tienen diferentes niveles de precisión o representatividad.

$$\bar{x}_p = \Sigma(w_i \times x_i) / \Sigma w_i$$

Figura 9. Fórmula.

Donde w_i = peso asignado a cada observación

Aplicación en producción animal:

Ejemplo práctico:

Una explotación ganadera tiene tres lotes de novillos con diferentes tamaños:

Tabla 20. Ganancia de peso en lotes bovinos.

Lote	Nº Animales	Ganancia Promedio (kg/día)
A	50	1.2
B	30	1.4
C	20	1.1

Media ponderada = $(50 \times 1.2 + 30 \times 1.4 + 20 \times 1.1) / 100 = 1.24$ kg/día

Interpretación: La ganancia promedio considerando el tamaño de cada lote es 1.24 kg/día, valor más representativo que la media simple (1.23 kg/día).

2.3.6 Implementación en R

```
# Ejemplo: Análisis de peso corporal en cachorros
# Datos de peso (kg) de 20 cachorros de 3 meses

# Cargar datos
pesos <- c(8.2, 7.5, 9.1, 8.8, 7.9, 8.5, 9.3, 8.1,
          7.7, 8.9, 8.4, 7.8, 9.0, 8.6, 8.3, 7.6,
          8.7, 9.2, 8.0, 8.5)

# Medidas de tendencia central
media_peso <- mean(pesos)
mediana_peso <- median(pesos)

# Crear tabla de frecuencias para identificar moda
tabla_freq <- table(pesos)
moda_peso <- as.numeric(names(tabla_freq)[which.max(tabla_freq)])

# Resultados
cat("Media:", round(media_peso, 2), "kg\n")
cat("Mediana:", mediana_peso, "kg\n")
cat("Moda:", moda_peso, "kg\n")

# Visualización comparativa
library(ggplot2)
df <- data.frame(peso = pesos)

ggplot(df, aes(x = peso)) +
  geom_histogram(bins = 8, fill = "lightblue", color = "black") +
  geom_vline(xintercept = media_peso, color = "red", linetype = "dashed", size = 1) +
  geom_vline(xintercept = mediana_peso, color = "blue", linetype = "dashed", size =
1) +
  annotate("text", x = media_peso + 0.2, y = 4, label = "Media", color = "red") +
  annotate("text", x = mediana_peso - 0.2, y = 4, label = "Mediana", color = "blue") +
```

Figura 10. Implementación en R.

2.4 Medidas de dispersión

2.4.1 Importancia de la variabilidad en medicina veterinaria

Las medidas de dispersión cuantifican la variabilidad presente en los datos, proporcionando información crucial sobre la homogeneidad de una población o la consistencia de un proceso (Nave et al., 2024). En veterinaria, comprender la dispersión resulta fundamental para:

- Establecer intervalos de referencia clínicos
- Evaluar la uniformidad en lotes de producción
- Detectar animales con valores anormales
- Diseñar experimentos con tamaños de muestra apropiados

2.4.2 Rango

El rango representa la diferencia entre el valor máximo y mínimo de un conjunto de datos.

$$\text{Rango} = X_{\text{máx}} - X_{\text{mín}}$$

Figura 11. Fórmula.

Limitaciones: Aunque simple de calcular, el rango solo considera valores extremos y puede ser altamente influenciado por valores atípicos.

Ejemplo práctico: Temperatura corporal en equinos

Temperaturas rectales (°C) en 10 caballos:

37.8, 38.2, 37.5, 38.0, 37.9, 38.5, 37.7, 38.1, 37.6, 38.3

Rango = 38.5 - 37.5 = 1.0°C

Interpretación: La variación total en temperatura es de 1°C, indicando relativa homogeneidad en el grupo estudiado.

2.4.3 Varianza

La varianza (s^2) mide la dispersión promedio de los datos respecto a la media, expresada en unidades al cuadrado.

$$s^2 = \Sigma(x_i - \bar{x})^2 / (n - 1)$$

Figura 12. Fórmula (muestra).

El denominador (n-1) representa los grados de libertad y proporciona un estimador insesgado de la varianza poblacional.

2.4.4 Desviación estándar

La desviación estándar (s) es la raíz cuadrada de la varianza, expresada en las mismas unidades que los datos originales.

$$s = \sqrt{s^2}$$

Figura 13. Fórmula.

Interpretación Práctica:

- Aproximadamente 68% de los datos se encuentran dentro de ± 1 desviación estándar de la media
- Aproximadamente 95% dentro de ± 2 desviaciones estándar
- Aproximadamente 99.7% dentro de ± 3 desviaciones estándar (para distribuciones normales)

Ejemplo práctico:

Niveles de cortisol ($\mu\text{g/dL}$) en 15 perros sometidos a situaciones estresantes:

```
r
# Datos de cortisol
cortisol <- c(8.2, 7.5, 9.8, 6.9, 10.2, 8.8, 7.1, 9.5,
             8.0, 11.3, 7.8, 9.1, 8.5, 10.5, 7.3)

# Cálculos estadísticos
media_cort <- mean(cortisol)
var_cort <- var(cortisol)
sd_cort <- sd(cortisol)

# Resultados
cat("Media:", round(media_cort, 2), " $\mu\text{g/dL}$ \n")
cat("Varianza:", round(var_cort, 2), " $(\mu\text{g/dL})^2$ \n")
cat("Desviación estándar:", round(sd_cort, 2), " $\mu\text{g/dL}$ \n")

# Intervalos de referencia (asumiendo normalidad)
lim_inf_1sd <- media_cort - sd_cort
lim_sup_1sd <- media_cort + sd_cort
lim_inf_2sd <- media_cort - 2*sd_cort
lim_sup_2sd <- media_cort + 2*sd_cort

cat("\nIntervalo  $\pm 1$  SD: [", round(lim_inf_1sd, 2), "-",
    round(lim_sup_1sd, 2), "]  $\mu\text{g/dL}$ \n")
cat("Intervalo  $\pm 2$  SD: [", round(lim_inf_2sd, 2), "-",
    round(lim_sup_2sd, 2), "]  $\mu\text{g/dL}$ \n")
```

Figura 14. Análisis de cortisol en perros con estrés.

Interpretación: Con media 8.63 µg/dL y desviación estándar 1.41 µg/dL, esperamos que aproximadamente 68% de los perros estresados presenten valores de cortisol entre 7.22 y 10.04 µg/dL.

2.4.5 Coeficiente de variación

El coeficiente de variación (CV) expresa la desviación estándar como porcentaje de la media, permitiendo comparar variabilidades entre variables con diferentes unidades o magnitudes.

$$CV = (s / \bar{x}) \times 100\%$$

Figura 15. Fórmula.

Interpretación en Veterinaria:

- CV < 10%: Baja variabilidad
- CV 10-20%: Variabilidad moderada
- CV > 20%: Alta variabilidad

Ejemplo práctico: Comparación de variabilidad en parámetros sanguíneos

Tabla 21. Análisis de dos parámetros en bovinos.

Parámetro	Media	Desviación Estándar	CV (%)
Hemoglobina (g/dL)	11.2	1.3	11.6

Leucocitos ($\times 10^3/\mu\text{L}$)	8.5	2.1	24.7
--	-----	-----	------

Interpretación: Aunque los leucocitos tienen menor desviación estándar absoluta, presentan mayor variabilidad relativa ($\text{CV}=24.7\%$) comparado con la hemoglobina ($\text{CV}=11.6\%$), indicando mayor heterogeneidad en el conteo leucocitario.

2.4.6 Rango intercuartílico

El rango intercuartílico (RIC) mide la dispersión del 50% central de los datos, siendo robusto ante valores atípicos.

$$\text{RIC} = Q3 - Q1$$

Figura 16. Fórmula.

Donde Q1 y Q3 son el primer y tercer cuartil, respectivamente.

Ejemplo práctico:

```

# Niveles de creatinina (mg/dL) en 30 gatos mayores de 10 años
creatinina <- c(1.2, 1.5, 1.8, 2.1, 1.4, 1.6, 1.9, 2.3, 1.7, 2.0,
               1.3, 1.8, 2.2, 1.5, 1.9, 2.4, 1.6, 2.1, 1.4, 1.8,
               2.5, 1.7, 2.0, 1.6, 3.8, 1.9, 2.2, 1.5, 4.2, 2.1)

# Cuartiles y RIC
Q1 <- quantile(creatinina, 0.25)
Q3 <- quantile(creatinina, 0.75)
RIC <- Q3 - Q1
mediana <- median(creatinina)

cat("Q1:", Q1, "mg/dL\n")
cat("Mediana:", mediana, "mg/dL\n")
cat("Q3:", Q3, "mg/dL\n")
cat("RIC:", RIC, "mg/dL\n")

# Identificación de valores atípicos (criterio 1.5×RIC)
lim_inf_outlier <- Q1 - 1.5 * RIC
lim_sup_outlier <- Q3 + 1.5 * RIC
outliers <- creatinina[creatinina < lim_inf_outlier |
                      creatinina > lim_sup_outlier]

cat("\nValores atípicos:", outliers, "mg/dL\n")

```

Figura 17. Análisis de creatinina en gatos geriátricos.

2.5 Medidas de forma y posición

2.5.1 Asimetría (Skewness)

La asimetría mide el grado de desviación de la simetría en una distribución. Esta medida resulta crucial para:

- Seleccionar pruebas estadísticas apropiadas
- Interpretar correctamente medidas de tendencia central
- Identificar patrones anormales en poblaciones animales

Coefficiente de Asimetría de Pearson:

$$As = 3(\bar{x} - Me) / s$$

Interpretación:

- $As = 0$: Distribución simétrica
- $As > 0$: Asimetría positiva (cola hacia la derecha)
- $As < 0$: Asimetría negativa (cola hacia la izquierda)

Ejemplo práctico:

```
# Edad (meses) al primer parto en 50 vaquillas
library(moments)

edad_parto <- c(24, 25, 23, 26, 24, 27, 25, 24, 28, 26,
  25, 24, 23, 29, 26, 25, 24, 30, 27, 25,
  24, 26, 25, 31, 24, 25, 27, 26, 24, 25,
  32, 25, 24, 26, 33, 25, 24, 27, 35, 26,
  24, 25, 36, 26, 24, 25, 38, 27, 24, 42)

# Cálculo de asimetría
asimetria <- skewness(edad_parto)
media_edad <- mean(edad_parto)
mediana_edad <- median(edad_parto)

cat("Media:", round(media_edad, 2), "meses\n")
cat("Mediana:", mediana_edad, "meses\n")
cat("Asimetría:", round(asimetria, 3), "\n")

# Visualización
library(ggplot2)
ggplot(data.frame(edad = edad_parto), aes(x = edad)) +
  geom_histogram(aes(y = ..density..), bins = 10,
    fill = "lightgreen", color = "black") +
  geom_density(color = "red", size = 1) +
  geom_vline(xintercept = media_edad, color = "blue",
    linetype = "dashed") +
  geom_vline(xintercept = mediana_edad, color = "purple",
    linetype = "dashed") +
  labs(title = "Distribución de Edad al Primer Parto",
    x = "Edad (meses)", y = "Densidad") +
  theme_minimal()
```

Figura 18. Distribución de edad en primera gestación en bovinos.

Interpretación: Con asimetría positiva (1.25), observamos que la mayoría de las vaquillas paren entre 23-27 meses, pero algunas presentan partos tardíos que extienden la cola hacia la derecha.

2.5.2 *Curtosis*

La curtosis mide el grado de concentración de los datos alrededor de la media, indicando si la distribución es más puntiaguda o achatada que una distribución normal.

Coefficiente de Curtosis:

$$K = \Sigma[(x_i - \bar{x})^4] / (n \times s^4) - 3$$

Interpretación:

- $K = 0$: Mesocúrtica (similar a normal)
- $K > 0$: Leptocúrtica (más puntiaguda)
- $K < 0$: Platicúrtica (más achatada)

Ejemplo práctico:

```
# Proteínas totales (g/dL) en gallinas ponedoras
proteinas <- c(3.2, 3.5, 3.8, 3.9, 3.9, 4.0, 4.0, 4.0, 4.1, 4.1,
              4.1, 4.1, 4.2, 4.2, 4.2, 4.2, 4.2, 4.3, 4.3, 4.3,
              4.3, 4.4, 4.4, 4.5, 4.8, 5.1, 5.4)

# Análisis de forma
kurt <- kurtosis(proteinas) - 3 # Exceso de curtosis
asim <- skewness(proteinas)

cat("Asimetría:", round(asim, 3), "\n")
cat("Exceso de curtosis:", round(kurt, 3), "\n")

# Test de normalidad Shapiro-Wilk
test_normal <- shapiro.test(proteinas)
cat("Test Shapiro-Wilk p-valor:", round(test_normal$p.value, 4), "\n")
```

Figura 19. Análisis de proteínas totales en aves.

2.5.3 Percentiles y cuartiles

Los percentiles dividen un conjunto de datos ordenados en 100 partes iguales, mientras que los cuartiles lo dividen en 4 partes.

Cuartiles especiales:

- Q1 (P25): 25% de los datos son menores
- Q2 (P50): Mediana
- Q3 (P75): 75% de los datos son menores

Aplicaciones veterinarias:

- Establecimiento de curvas de crecimiento
- Determinación de rangos de referencia clínicos
- Clasificación de condición corporal

Ejemplo práctico:

```
# Peso (kg) de 100 lechones a los 21 días
set.seed(123)
pesos_lechones <- rnorm(100, mean = 6.5, sd = 0.8)

# Cálculo de percentiles
percentiles <- quantile(pesos_lechones,
  probs = c(0.05, 0.10, 0.25, 0.50,
    0.75, 0.90, 0.95))

# Crear tabla de referencia
tabla_percentiles <- data.frame(
  Percentil = c("P5", "P10", "P25", "P50", "P75", "P90", "P95"),
  Peso_kg = round(percentiles, 2)
)

print(tabla_percentiles)

# Clasificación de lechones
clasificar_lechon <- function(peso) {
  if(peso < percentiles[2]) return("Bajo peso")
  else if(peso < percentiles[5]) return("Normal")
  else if(peso < percentiles[6]) return("Sobre promedio")
  else return("Peso excepcional")
}

# Ejemplo de clasificación
peso_prueba <- 7.2
cat("\nLechón con", peso_prueba, "kg:",
  clasificar_lechon(peso_prueba), "\n")
```

Figura 20. Curva de crecimiento en lechones.

2.6 Representaciones gráficas

2.6.1 Principios de visualización de datos veterinarios

La representación gráfica efectiva de datos veterinarios debe:

- Comunicar información de manera clara y precisa

- Revelar patrones y tendencias relevantes
- Facilitar la comparación entre grupos
- Identificar valores atípicos o anormales
- Respetar las propiedades de las variables representadas

2.6.2 Histogramas

Los histogramas representan la distribución de frecuencias de variables continuas mediante barras contiguas.

Elementos Clave:

- Número de clases (regla de Sturges): $k \approx 1 + 3.322 \times \log(n)$
- Amplitud de clase constante
- Área proporcional a la frecuencia

Ejemplo práctico:

```
# Pesos (kg) de 150 cerdos a los 150 días
set.seed(456)
pesos_cerdos <- rnorm(150, mean = 95, sd = 8)

# Histograma con curva normal superpuesta
library(ggplot2)
ggplot(data.frame(peso = pesos_cerdos), aes(x = peso)) +
  geom_histogram(aes(y = ..density..), bins = 12,
    fill = "coral", color = "black", alpha = 0.7) +
  stat_function(fun = dnorm,
    args = list(mean = mean(pesos_cerdos),
      sd = sd(pesos_cerdos)),
    color = "darkblue", size = 1.2) +
  labs(title = "Distribución de Pesos en Cerdos de Engorde (150
    días)",
    x = "Peso (kg)", y = "Densidad") +
  theme_classic() +
  annotate("text", x = 110, y = 0.04,
    label = paste("n =", length(pesos_cerdos), "\n",
      "Media =", round(mean(pesos_cerdos), 1), "kg\n",
      "SD =", round(sd(pesos_cerdos), 1), "kg"))
```

Figura 21. Distribución de pesos en cerdos de engorde.

2.6.3 Diagramas de caja (Box Plots)

Los diagramas de caja resumen cinco estadísticos fundamentales y permiten identificar valores atípicos.

Componentes:

- Caja: Q1 a Q3 (50% central de los datos)
- Línea central: Mediana
- Bigotes: Valores extremos dentro de $1.5 \times \text{RIC}$
- Puntos: Valores atípicos

Ejemplo práctico:

```
# Niveles de ALT (UI/L) en tres razas
set.seed(789)
labrador <- rnorm(30, mean = 45, sd = 12)
pastor_aleman <- rnorm(30, mean = 52, sd = 15)
bulldog <- rnorm(30, mean = 58, sd = 18)

# Crear dataframe
alt_data <- data.frame(
  ALT = c(labrador, pastor_aleman, bulldog),
  Raza = factor(rep(c("Labrador", "Pastor Alemán", "Bulldog"),
                    each = 30))
)

# Box plot comparativo
ggplot(alt_data, aes(x = Raza, y = ALT, fill = Raza)) +
  geom_boxplot(alpha = 0.7, outlier.color = "red",
               outlier.shape = 16, outlier.size = 2) +
  geom_jitter(width = 0.2, alpha = 0.3, size = 1) +
  scale_fill_manual(values = c("lightblue", "lightgreen", "lightyellow")) +
  labs(title = "Comparación de Niveles de ALT entre Razas Caninas",
       y = "ALT (UI/L)") +
  theme_minimal() +
  theme(legend.position = "none") +
  # Línea de referencia para valor normal máximo
  geom_hline(yintercept = 70, linetype = "dashed", color = "red") +
  annotate("text", x = 3.3, y = 72, label = "Límite superior normal",
         color = "red", size = 3)
```

Figura 22. Comparación de ALT entre razas caninas.

2.6.4 Diagramas de dispersión

Los diagramas de dispersión visualizan la relación entre dos variables continuas.

Ejemplo práctico:

```
# Datos de 40 potros
set.seed(321)
edad_meses <- sample(1:24, 40, replace = TRUE)
peso_kg <- 50 + 15 * edad_meses + rnorm(40, mean = 0, sd = 10)

datos_potros <- data.frame(edad = edad_meses, peso = peso_kg)

# Diagrama de dispersión con línea de tendencia
ggplot(datos_potros, aes(x = edad, y = peso)) +
  geom_point(size = 3, color = "darkgreen", alpha = 0.6) +
  geom_smooth(method = "lm", se = TRUE, color = "blue") +
  labs(title = "Relación Peso-Edad en Potros",
        x = "Edad (meses)", y = "Peso (kg)") +
  theme_classic() +
  # Añadir ecuación de regresión
  stat_poly_eq(formula = y ~ x,
               aes(label = paste(..eq.label.., ..rr.label..,
                                sep = "~~~")),
               parse = TRUE, size = 4)
```

Figura 23. Relación peso-edad en potros.

2.6.5 Gráficos de barras y columnas

Ideales para variables categóricas o comparaciones entre grupos

Ejemplo práctico:

```

# Casos diagnosticados en un trimestre
enfermedades <- data.frame(
  Patologia = c("Dermatitis", "Otitis", "Gastroenteritis",
    "Enfermedad periodontal", "Obesidad", "Artritis",
    "Diabetes", "Insuficiencia renal"),
  Casos = c(145, 98, 87, 76, 65, 43, 28, 15),
  Porcentaje = c(26.0, 17.6, 15.6, 13.6, 11.7, 7.7, 5.0, 2.7)
)

# Ordenar por frecuencia
enfermedades <- enfermedades[order(enfermedades$Casos, decreasing = TRUE),]

# Gráfico de barras horizontal
library(ggplot2)
ggplot(enfermedades, aes(x = reorder(Patologia, Casos), y = Casos)) +
  geom_bar(stat = "identity", fill = "steelblue", alpha = 0.8) +
  geom_text(aes(label = paste0(Casos, " (", Porcentaje, "%)"),
    hjust = -0.1, size = 3)) +
  coord_flip() +
  labs(title = "Prevalencia de Enfermedades en Pequeños Animales",
    subtitle = "Datos trimestrales de clínica veterinaria",
    x = "", y = "Número de casos") +
  theme_minimal() +
  theme(panel.grid.major.y = element_blank()) +
  scale_y_continuous(expand = expansion(mult = c(0, 0.15)))

```

Figura 24. Prevalencia de enfermedades en una clínica veterinaria.

2.6.6 Gráficos circulares (Pie Charts)

Aunque menos recomendados que los gráficos de barras para comparaciones precisas, los gráficos circulares pueden ser útiles para mostrar proporciones del total.

Ejemplo práctico:

```

# Proporción de especies en hospital veterinario
especies <- data.frame(
  Especie = c("Caninos", "Felinos", "Aves", "Roedores", "Reptiles", "Otros"),
  Cantidad = c(520, 380, 85, 45, 25, 20)
)

# Calcular porcentajes
especies$Porcentaje <- round(especies$Cantidad / sum(especies$Cantidad) *
100, 1)

# Gráfico circular mejorado
library(ggplot2)
library(scales)

ggplot(especies, aes(x = "", y = Cantidad, fill = Especie)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y", start = 0) +
  scale_fill_brewer(palette = "Set3") +
  geom_text(aes(label = paste0(Porcentaje, "%")),
    position = position_stack(vjust = 0.5)) +
  labs(title = "Distribución de Especies Atendidas",
    subtitle = "Hospital Veterinario - Año 2024") +
  theme_void() +
  theme(legend.position = "right")

```

Figura 25. Distribución de especies atendidas.

2.6.7 Gráficos de líneas

Fundamentales para representar series temporales y tendencias.

Ejemplo práctico:

```

# Producción mensual de leche (L/vaca/día)
meses <- 1:12
produccion <- c(28.5, 29.2, 30.8, 32.1, 33.5, 34.2,
               33.8, 32.5, 30.9, 29.1, 27.8, 27.2)
temperatura <- c(15, 17, 20, 23, 26, 29, 31, 30, 27, 23, 19, 16)

datos_produccion <- data.frame(
  Mes = factor(month.abb, levels = month.abb),
  Produccion = produccion,
  Temperatura = temperatura
)

# Gráfico de doble eje
library(ggplot2)
coeff <- 0.5 # Coeficiente para escalar el segundo eje

ggplot(datos_produccion, aes(x = Mes)) +

```

Figura 26. Evolución de producción láctea.

2.6.8 Mapas de calor (Heatmaps)

Útiles para visualizar matrices de correlación o datos multidimensionales.

Ejemplo práctico:

```

# Generar datos de parámetros sanguíneos
set.seed(654)
n <- 100
parametros_sangre <- data.frame(
  Hemoglobina = rnorm(n, 14, 1.5),
  Hematocrito = rnorm(n, 42, 4),
  Leucocitos = rnorm(n, 8000, 1500),
  Plaquetas = rnorm(n, 250000, 50000),
  Proteinas = rnorm(n, 7, 0.8),
  Albumina = rnorm(n, 3.5, 0.4)
)

# Matriz de correlación
cor_matrix <- cor(parametros_sangre)

# Preparar datos para heatmap
library(reshape2)
cor_melted <- melt(cor_matrix)

# Heatmap
ggplot(cor_melted, aes(x = Var1, y = Var2, fill = value)) +
  geom_tile() +
  geom_text(aes(label = round(value, 2)), color = "white", size = 4) +
  scale_fill_gradient2(low = "blue", high = "red", mid = "white",
    midpoint = 0, limit = c(-1, 1),
    name = "Correlación") +
  labs(title = "Matriz de Correlación - Parámetros Hematológicos",
    x = "", y = "") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

```

Figura 27. Correlación entre parámetros hematológicos.

2.7 Aplicaciones en parámetros clínicos veterinarios

2.7.1 Establecimiento de valores de referencia

El establecimiento de valores de referencia constituye una aplicación fundamental de la estadística descriptiva en medicina veterinaria. Estos valores proporcionan la base para la interpretación de resultados diagnósticos (Taylor & Meyer, 2023).

Metodología para establecer intervalos de referencia:

1. **Selección de población de referencia:** Animales clínicamente sanos
2. **Tamaño de muestra mínimo:** $n \geq 120$ (recomendación CLSI)
3. **Análisis de distribución:** Verificar normalidad
4. **Cálculo del intervalo:** Percentiles 2.5 y 97.5

Ejemplo práctico:

```

# Simulación de 150 valores de glucosa en gatos sanos
set.seed(987)
glucosa_gatos <- rnorm(150, mean = 95, sd = 15)

# Eliminar valores extremos (outliers)
Q1 <- quantile(glucosa_gatos, 0.25)
Q3 <- quantile(glucosa_gatos, 0.75)
IQR <- Q3 - Q1
outliers_inf <- Q1 - 3 * IQR
outliers_sup <- Q3 + 3 * IQR

glucosa_clean <- glucosa_gatos[glucosa_gatos >= outliers_inf &
  glucosa_gatos <= outliers_sup]

# Test de normalidad
shapiro_test <- shapiro.test(glucosa_clean)

if(shapiro_test$p.value > 0.05) {
  # Método paramétrico
  ref_inf <- mean(glucosa_clean) - 1.96 * sd(glucosa_clean)
  ref_sup <- mean(glucosa_clean) + 1.96 * sd(glucosa_clean)
  metodo <- "Paramétrico"
} else {
  # Método no paramétrico
  ref_inf <- quantile(glucosa_clean, 0.025)
  ref_sup <- quantile(glucosa_clean, 0.975)
  metodo <- "No paramétrico"
}

cat("Intervalo de Referencia para Glucosa en Felinos\n")
cat("Método:", metodo, "\n")
cat("n =", length(glucosa_clean), "gatos\n")
cat("Intervalo 95%: [", round(ref_inf, 1), "-",
  round(ref_sup, 1), "] mg/dL\n")

# Visualización
library(ggplot2)
ggplot(data.frame(glucosa = glucosa_clean), aes(x = glucosa)) +
  geom_histogram(aes(y = ..density..), bins = 15,
    fill = "lightblue", color = "black", alpha = 0.7) +
  geom_density(color = "red", size = 1) +
  geom_vline(xintercept = c(ref_inf, ref_sup),
    color = "darkgreen", linetype = "dashed", size = 1) +
  annotate("rect", xmin = ref_inf, xmax = ref_sup,
    ymin = 0, ymax = Inf, alpha = 0.2, fill = "green") +
  labs(title = "Distribución de Glucosa Sanguínea en Felinos Sanos",
    subtitle = paste("Intervalo de referencia 95%:",
      round(ref_inf, 1), "-", round(ref_sup, 1), "mg/dL"),
    x = "Glucosa (mg/dL)", y = "Densidad") +
  theme_classic()

```

Figura 28. Valores de referencia para glucosa en felinos.

2.7.2 Análisis de parámetros reproductivos

Ejemplo práctico:

```

# Intervalos entre partos (días) en 80 vacas lecheras
set.seed(111)
intervalos_parto <- c(
  rnorm(60, mean = 385, sd = 25), # Mayoría con intervalos normales
  rnorm(15, mean = 450, sd = 30), # Algunas con intervalos largos
  rnorm(5, mean = 340, sd = 15) # Pocas con intervalos cortos)
# Análisis descriptivo completo
analisis_reproductivo <- function(datos) {
  resultado <- list(
    n = length(datos),
    media = mean(datos),
    mediana = median(datos),
    sd = sd(datos),
    cv = (sd(datos)/mean(datos)) * 100,
    min = min(datos),
    max = max(datos),
    Q1 = quantile(datos, 0.25),
    Q3 = quantile(datos, 0.75),
    IQR = IQR(datos),
    asimetria = moments::skewness(datos),
    curtosis = moments::kurtosis(datos) - 3 )
  # Clasificación por categorías
  resultado$optimo <- sum(datos >= 365 & datos <= 395)
  resultado$aceptable <- sum(datos > 395 & datos <= 420)
  resultado$largo <- sum(datos > 420)
  resultado$corto <- sum(datos < 365)
  return(resultado)}
res <- analisis_reproductivo(intervalos_parto)
# Imprimir resultados
cat("ANÁLISIS DE INTERVALOS ENTRE PARTOS\n")
cat("=====\n")
cat("Estadísticos descriptivos:\n")
cat(" N:", res$n, "vacas\n")
cat(" Media:", round(res$media, 1), "días\n")
cat(" Mediana:", round(res$mediana, 1), "días\n")
cat(" Desviación estándar:", round(res$sd, 1), "días\n")
cat(" CV:", round(res$cv, 1), "%\n")
cat(" Rango: [", round(res$min, 0), "-", round(res$max, 0), "] días\n\n")

cat("Clasificación reproductiva:\n")
cat(" Óptimo (365-395 días):", res$optimo, "vacas (",
  round(res$optimo/res$n*100, 1), "%)\n")
cat(" Aceptable (396-420 días):", res$aceptable, "vacas (",
  round(res$aceptable/res$n*100, 1), "%)\n")
cat(" Largo (>420 días):", res$largo, "vacas (",
  round(res$largo/res$n*100, 1), "%)\n")
cat(" Corto (<365 días):", res$corto, "vacas (",
  round(res$corto/res$n*100, 1), "%)\n")

```

Figura 29. Análisis descriptivo de intervalos entre partos en bovinos.

2.7.3 Monitoreo de crecimiento animal

Ejemplo práctico:

```

# Datos de peso semanal en pollos (g)
semanas <- 1:6
n_pollos <- 100

# Simular datos de crecimiento
set.seed(222)
datos_crecimiento <- data.frame()
for(semana in semanas) {
  peso_esperado <- 180 * semana + 50 * semana^1.8
  pesos <- rnorm(n_pollos, mean = peso_esperado, sd = peso_esperado * 0.08)
  temp_df <- data.frame(
    Semana = semana,
    Peso = pesos )
  datos_crecimiento <- rbind(datos_crecimiento, temp_df)}
# Calcular estadísticos por semana
library(dplyr)
resumen_semanal <- datos_crecimiento %>%
  group_by(Semana) %>%
  summarise(
    Media = mean(Peso),
    SD = sd(Peso),
    CV = (SD/Media) * 100,
    P10 = quantile(Peso, 0.10),
    P50 = quantile(Peso, 0.50),
    P90 = quantile(Peso, 0.90),
    .groups = 'drop' )
# Gráfico de curvas de crecimiento
ggplot(resumen_semanal, aes(x = Semana)) +
  geom_ribbon(aes(ymin = P10, ymax = P90), fill = "gray80", alpha = 0.5) +
  geom_line(aes(y = Media), color = "red", size = 2) +
  geom_line(aes(y = P10), color = "blue", linetype = "dashed") +
  geom_line(aes(y = P90), color = "blue", linetype = "dashed") +
  geom_point(aes(y = Media), size = 3, color = "red") +
  scale_x_continuous(breaks = 1:6) +
  labs(title = "Curva de Crecimiento en Pollos de Engorde",
    subtitle = "Media y percentiles 10-90",
    x = "Edad (semanas)",
    y = "Peso corporal (g)") +
  theme_classic() +
  annotate("text", x = 5.5, y = resumen_semanal$Media[6],
    label = "Media", color = "red") +
  annotate("text", x = 5.5, y = resumen_semanal$P90[6],
    label = "P90", color = "blue") +
  annotate("text", x = 5.5, y = resumen_semanal$P10[6],
    label = "P10", color = "blue")
# Tabla de referencia
print(resumen_semanal)

```

Figura 30. Curvas de crecimiento en pollos de engorde.

2.7.4 Control de calidad en laboratorio veterinario

Ejemplo práctico:

```

# Control de calidad para análisis de hemoglobina
# 30 días de mediciones del control
set.seed(333) dias <- 1:30
valor_target <- 12.0 # g/dL
sd_metodo <- 0.3
# Simular mediciones con deriva gradual
mediciones_control <- valor_target + rnorm(30, mean = 0, sd = sd_metodo)
+      seq(0, 0.3, length.out = 30) # Deriva sistemática
# Calcular límites de control
media_control <- valor_target
limite_2sd_sup <- media_control + 2 * sd_metodo
limite_2sd_inf <- media_control - 2 * sd_metodo
limite_3sd_sup <- media_control + 3 * sd_metodo
limite_3sd_inf <- media_control - 3 * sd_metodo
# Identificar violaciones de reglas de Westgard
violaciones <- character()
if(any(mediciones_control > limite_3sd_sup | mediciones_control <
limite_3sd_inf)) { violaciones <- c(violaciones, "1-3s: Valor fuera de
±3SD")} if(sum(mediciones_control > limite_2sd_sup) >= 2) { violaciones
<- c(violaciones, "2-2s: Dos valores consecutivos > 2SD")}
}# Gráfico de Levey-Jennings df_control <- data.frame(Dia = dias,
Hemoglobina = mediciones_control)
ggplot(df_control, aes(x = Dia, y = Hemoglobina)) +
  geom_line(color = "black", size = 1) +
  geom_point(size = 2, color = ifelse(abs(mediciones_control -
media_control) > 2*sd_metodo,
      "red", "black")) + geom_hline(yintercept =
media_control, color = "green", size = 1) +
  geom_hline(yintercept = c(limite_2sd_sup, limite_2sd_inf),
      color = "orange", linetype = "dashed") +
  geom_hline(yintercept = c(limite_3sd_sup, limite_3sd_inf),
      color = "red", linetype = "dashed") +
  annotate("text", x = 29, y = media_control, label = "Media", color =
"green") +
  annotate("text", x = 29, y = limite_2sd_sup, label = "+2SD", color =
"orange") +
  annotate("text", x = 29, y = limite_2sd_inf, label = "-2SD", color =
"orange") +
  annotate("text", x = 29, y = limite_3sd_sup, label = "+3SD", color = "red")
+ labs(title = "Gráfico de Control de Levey-Jennings",
      subtitle = "Control de Hemoglobina - 30 días",
      x = "Día", y = "Hemoglobina (g/dL)") + theme_classic()
# Imprimir violaciones detectadas
if(length(violaciones) > 0) { cat("ALERTAS DE CONTROL DE
CALIDAD:\n") for(v in violaciones) cat("-", v, "\n")}

```

Figura 31. Gráficos de control de Levey-Jennings.

2.8 Conclusiones del capítulo

La estadística descriptiva constituye una herramienta fundamental en el análisis inicial de los datos veterinarios, ya que permite una caracterización precisa y ordenada de la información recolectada. En este capítulo se ha demostrado que las medidas de tendencia central son indispensables para identificar valores típicos, siempre considerando la necesidad de seleccionar la medida más adecuada según la distribución de los datos y la presencia de valores atípicos. Del mismo modo, las medidas de dispersión resultan esenciales para evaluar la variabilidad biológica, establecer intervalos de referencia confiables y detectar posibles problemas en procesos productivos o diagnósticos clínicos.

Por su parte, las medidas de forma y posición aportan una comprensión más profunda de la distribución de los datos, facilitando la identificación de desviaciones de la normalidad que condicionan la elección de pruebas estadísticas posteriores. Asimismo, las representaciones gráficas se consolidan como herramientas poderosas para comunicar información compleja de manera clara y efectiva, permitiendo identificar patrones, tendencias y anomalías relevantes. La aplicación práctica de estos conceptos en parámetros clínicos evidencia su valor en la toma de decisiones veterinarias fundamentadas en evidencia.

En conjunto, el dominio de estas técnicas descriptivas sienta las bases para análisis inferenciales más complejos, posibilitando una correcta interpretación de los datos en la práctica clínica, la investigación y la producción animal. Finalmente, la implementación de estas

herramientas mediante software estadístico como R optimiza el manejo de grandes volúmenes de datos y favorece una medicina veterinaria más precisa, confiable y sustentada en evidencia cuantitativa.

CAPÍTULO III

3 PROBABILIDAD Y DISTRIBUCIONES

3.1 Objetivos de aprendizaje

Al finalizar este capítulo, el estudiante será capaz de:

1. Comprender y aplicar los conceptos fundamentales de probabilidad en contextos veterinarios
2. Identificar y utilizar las distribuciones de probabilidad más relevantes en medicina veterinaria
3. Aplicar el teorema del límite central en investigaciones con poblaciones animales
4. Resolver problemas prácticos relacionados con probabilidades en diagnóstico y epidemiología veterinaria
5. Interpretar distribuciones muestrales para la toma de decisiones clínicas informadas

3.2 Conceptos básicos de probabilidad

3.2.1 Introducción a la probabilidad en ciencias veterinarias

La probabilidad constituye el fundamento matemático para la toma de decisiones en medicina veterinaria. Desde el diagnóstico clínico hasta el diseño de programas de control epidemiológico, el profesional veterinario enfrenta constantemente situaciones de incertidumbre que requieren un análisis probabilístico riguroso (Kaps & Lamberson, 2004). La comprensión de estos conceptos permite al veterinario

cuantificar la incertidumbre inherente a los procesos biológicos y tomar decisiones basadas en evidencia.

3.2.2 Definiciones fundamentales

Experimento aleatorio: Un procedimiento que puede repetirse bajo las mismas condiciones y cuyo resultado no puede predecirse con certeza. En el contexto veterinario, la administración de un tratamiento antiparasitario representa un experimento aleatorio, donde la respuesta individual de cada animal puede variar.

Espacio muestral (S): El conjunto de todos los resultados posibles de un experimento aleatorio. Por ejemplo, al evaluar la condición corporal de bovinos en una escala del 1 al 5, el espacio muestral sería $S = \{1, 2, 3, 4, 5\}$.

Evento: Un subconjunto del espacio muestral. En un estudio de prevalencia de brucelosis bovina, el evento A podría definirse como "animal seropositivo".

3.2.3 Enfoques de la probabilidad

Enfoque clásico (A Priori)

La probabilidad de un evento A se calcula como:

$$P(A) = \frac{\text{Número de casos favorables}}{\text{Número total de casos posibles}}$$

Este enfoque asume que todos los resultados son igualmente probables. En genética veterinaria, al cruzar dos heterocigotos ($Aa \times Aa$), la probabilidad de obtener un homocigoto recesivo (aa) sería $P(aa) = 1/4$.

Enfoque Frecuentista

Basado en la repetición del experimento un gran número de veces:

$$P(A) = \lim_{n \rightarrow \infty} \frac{f_A}{n}$$

Donde f_A representa la frecuencia absoluta del evento A y n el número total de observaciones. Este enfoque resulta particularmente útil en estudios epidemiológicos longitudinales.

Enfoque Subjetivo

Representa el grado de creencia basado en la experiencia y conocimiento del observador. Un clínico veterinario experimentado puede asignar una probabilidad subjetiva al diagnóstico diferencial basándose en la anamnesis y signos clínicos observados (Lamb, 2016).

3.2.4 Axiomas de Probabilidad

Los axiomas de Kolmogorov establecen las bases matemáticas de la teoría de probabilidades:

- **No negatividad:** $P(A) \geq 0$ para cualquier evento A
- **Normalización:** $P(S) = 1$
- **Aditividad:** Para eventos mutuamente excluyentes A y B: $P(A \cup B) = P(A) + P(B)$

3.2.5 Reglas fundamentales de probabilidad

a) Regla de la adición

Para eventos no mutuamente excluyentes:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Ejemplo veterinario: En una población canina, la probabilidad de encontrar un perro con sobrepeso es 0.40, la probabilidad de diabetes es 0.15, y la probabilidad de ambas condiciones es 0.08. La probabilidad de encontrar un perro con al menos una de estas condiciones sería:

$$P(\text{Sobrepeso} \cup \text{Diabetes}) = 0.40 + 0.15 - 0.08 = 0.47$$

Probabilidad Condicional

La probabilidad de que ocurra A dado que B ha ocurrido:

$$P(A | B) = \frac{P(A \cap B)}{P(B)}, \text{ donde } P(B) > 0$$

Esta conceptualización resulta fundamental en el diagnóstico veterinario. Por ejemplo, la probabilidad de que un gato presente insuficiencia renal crónica dado que presenta poliuria/polidipsia.

b) Regla de la multiplicación

$$P(A \cap B) = P(A/B) \cdot P(B) = P(B/A) \cdot P(A)$$

Independencia Estadística

Dos eventos son independientes si:

- $P(A|B) = P(A)$
- $P(B|A) = P(B)$
- $P(A \cap B) = P(A) \cdot P(B)$

3.2.6 Teorema de bayes en medicina veterinaria

Como lo menciona (Watty, 2016) el teorema de Bayes permite actualizar probabilidades a priori con nueva información:

$$P(A | B) = \frac{P(B | A) \cdot P(A)}{P(B)}$$

Expandiendo el denominador mediante la ley de probabilidad total:

$$P(A | B) = \frac{P(B | A) \cdot P(A)}{P(B | A) \cdot P(A) + P(B | A^c) \cdot P(A^c)}$$

Aplicación en Diagnóstico Veterinario

Consideremos una prueba diagnóstica para leptospirosis canina con:

- Sensibilidad = 0.95 (probabilidad de resultado positivo dado que el animal está enfermo)
- Especificidad = 0.90 (probabilidad de resultado negativo dado que el animal está sano)
- Prevalencia = 0.10 (probabilidad a priori de la enfermedad)

El valor predictivo positivo (VPP) se calcula:

$$VPP = \frac{0.95 \times 0.10}{(0.95 \times 0.10) + (0.10 \times 0.90)} = \frac{0.095}{0.185} = 0.514$$

Este resultado indica que solo el 51.4% de los animales con prueba positiva realmente tienen la enfermedad, destacando la importancia de considerar la prevalencia en la interpretación de pruebas diagnósticas.

3.3 Distribución normal

3.3.1 Características y propiedades

La distribución normal o gaussiana representa el modelo probabilístico más importante en bioestadística veterinaria. Numerosas variables biológicas siguen aproximadamente esta distribución: peso corporal, niveles de glucosa sanguínea, producción láctea, entre otras (Anderson et al., 2017).

La función de densidad de probabilidad está dada por:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Donde:

- μ = media poblacional
- σ = desviación estándar poblacional
- $\pi \approx 3.14159$
- $e \approx 2.71828$

Propiedades fundamentales:

1. **Simetría:** La distribución es perfectamente simétrica respecto a la media
2. **Unimodalidad:** Presenta un único máximo en $x = \mu$
3. **Asíntota horizontal:** Las colas se aproximan al eje x pero nunca lo tocan
4. **Área total:** El área bajo la curva es igual a 1
5. **Regla empírica (68-95-99.7):**
 - $P(\mu - \sigma < X < \mu + \sigma) \approx 0.68$
 - $P(\mu - 2\sigma < X < \mu + 2\sigma) \approx 0.95$
 - $P(\mu - 3\sigma < X < \mu + 3\sigma) \approx 0.997$

3.3.2 Distribución normal estándar

La estandarización transforma cualquier distribución normal en una distribución normal estándar con $\mu = 0$ y $\sigma = 1$:

$$Z = \frac{X - \mu}{\sigma}$$

Esta transformación facilita el uso de tablas estadísticas y la comparación entre diferentes variables.

Ejemplo práctico: El peso al nacimiento de terneros Holstein sigue una distribución normal con $\mu = 42$ kg y $\sigma = 5$ kg. Para calcular la probabilidad de que un ternero pese más de 48 kg:

$$Z = \frac{48 - 42}{5} = 1.2$$

Consultando la tabla normal estándar: $P(Z > 1.2) = 0.1151$

Por lo tanto, aproximadamente el 11.51% de los terneros pesarán más de 48 kg al nacer.

3.3.3 Aplicaciones en medicina veterinaria

Intervalos de referencia clínicos

Los valores de referencia para parámetros hematológicos y bioquímicos se establecen típicamente como $\mu \pm 2\sigma$, abarcando aproximadamente el 95% de la población sana. Por ejemplo, para la concentración de hemoglobina en perros adultos sanos ($\mu = 15.5$ g/dL, $\sigma = 1.8$ g/dL), el intervalo de referencia sería:

Límite inferior = $15.5 - (2 \times 1.8) = 11.9$ g/dL Límite superior = $15.5 + (2 \times 1.8) = 19.1$ g/dL

Control de calidad en producción animal

En una granja avícola, el peso de pollos de engorde a las 6 semanas sigue $N(2.5 \text{ kg}, 0.3 \text{ kg})$. Para mantener estándares de calidad, se rechazan aves con peso inferior a 2 kg o superior a 3 kg.

Proporción rechazada por bajo peso:

$$P(X < 2) = P\left(Z \frac{2 - 2.5}{0.3}\right) = P(Z < -1.67) = 0.0475$$

Proporción rechazada por exceso de peso:

$$P(X < 3) = P\left(Z \frac{3 - 2.5}{0.3}\right) = P(Z < -1.67) = 0.0475$$

Proporción total rechazada = $0.0475 + 0.0475 = 0.095$ (9.5%)

3.3.4 Verificación de normalidad

Antes de aplicar métodos paramétricos, es fundamental verificar el supuesto de normalidad mediante:

Métodos gráficos:

- Histogramas con curva normal superpuesta
- Gráficos Q-Q (cuantil-cuantil)
- Gráficos de probabilidad normal

Pruebas estadísticas:

- Shapiro-Wilk ($n < 50$)
- Kolmogorov-Smirnov con corrección de Lilliefors ($n \geq 50$)
- Anderson-Darling

3.4 Distribución Binomial y de Poisson

3.4.1 Distribución Binomial

La distribución binomial modela el número de éxitos en n ensayos independientes de Bernoulli, cada uno con probabilidad de éxito p .

Función de probabilidad:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Donde

$$\binom{n}{k} = \frac{n!}{k! (n - k)!}$$

Parámetros:

- Media: $\mu = np$
- Varianza: $\sigma^2 = np(1-p)$
- Desviación estándar: $\sigma = \sqrt{np(1-p)}$

Condiciones de aplicación:

1. Número fijo de ensayos (n)
2. Cada ensayo tiene dos resultados posibles (éxito/fracaso)
3. Probabilidad de éxito (p) constante
4. Ensayos independientes

Ejemplo veterinario - Eficacia vacunal:

Una vacuna contra parvovirus canino tiene una eficacia del 85%. Si se vacunan 20 cachorros, ¿cuál es la probabilidad de que exactamente 18 desarrollen inmunidad?

$$P(X = 18) = \frac{20}{18} (0.85)^{18} (0.15)^2$$

$$P(X = 18) = 190 \times 0.0388 \times 0.0225 = 0.166$$

3.4.2 Distribución de Poisson

La distribución de Poisson modela el número de eventos que ocurren en un intervalo de tiempo o espacio fijo, cuando estos eventos ocurren con una tasa media constante λ .

Función de probabilidad:

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

Parámetros:

- Media: $\mu = \lambda$
- Varianza: $\sigma^2 = \lambda$
- Desviación estándar: $\sigma = \sqrt{\lambda}$

Aplicación en Epidemiología Veterinaria:

En una clínica veterinaria se atienden en promedio 3 emergencias por día ($\lambda = 3$). La probabilidad de atender exactamente 5 emergencias en un día es:

$$P(X = 5) = \frac{e^{-3} * 3^5}{5!} = \frac{0.0498 \times 243}{120} = 0.101$$

Aproximación de Binomial a Poisson:

Cuando n es grande ($n > 30$) y p es pequeño ($p < 0.1$), la distribución binomial puede aproximarse mediante Poisson con $\lambda = np$.

3.4.3 Comparación y selección de distribuciones

Tabla 22. Comparación y selección de distribuciones.

Característica	Binomial	Poisson
Tipo de variable	Número de éxitos en n ensayos	Número de eventos en intervalo fijo
Parámetros	n, p	λ
Media	np	λ
Varianza	$np(1-p)$	λ
Aplicación veterinaria típica	Prevalencia, eficacia de tratamiento	Incidencia, casos raros

3.5 Distribuciones muestrales

3.5.1 Concepto de distribución muestral

Una distribución muestral describe la variabilidad de un estadístico calculado a partir de todas las posibles muestras de tamaño n extraídas de una población. Este concepto resulta fundamental para la inferencia estadística en investigación veterinaria (Jiang, 2022).

3.5.2 Distribución muestral de la media

Para una población con media μ y desviación estándar σ , la distribución muestral de la media muestral ($\bar{X}$) tiene las siguientes propiedades:

Media de la distribución muestral:

$$\mu_{\bar{X}} = \mu$$

Error estándar de la media:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

El error estándar disminuye con el tamaño de muestra, reflejando mayor precisión en la estimación.

Ejemplo práctico - producción láctea:

En un hato lechero, la producción diaria de leche sigue $N(25 \text{ L}, 4 \text{ L})$. Si se seleccionan aleatoriamente 16 vacas:

- Media de la distribución muestral: $\mu_{\bar{X}} = 25 \text{ L}$
- Error estándar: $\sigma_{\bar{X}} = 4/\sqrt{16} = 1 \text{ L}$

La probabilidad de que la producción promedio de la muestra supere los 26 L:

$$Z = \frac{26 - 25}{1} = 1$$

$$P(\bar{X} > 26) = P(Z > 1) = 0.1587$$

3.5.3 Distribución muestral de la proporción

Para una proporción poblacional p , la distribución muestral de la proporción muestral ($\hat{p}$) tiene:

Media:

$$\mu_{\hat{p}} = p$$

Error estándar:

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

Condiciones para aproximación normal:

- $np \geq 5$
- $n(1-p) \geq 5$

Aplicación en estudios de prevalencia:

En un estudio de seroprevalencia de brucelosis bovina con prevalencia esperada del 8%, si se muestrean 200 animales:

- Error estándar: $\sigma_{\hat{p}} = \sqrt{[(0.08 \times 0.92)/200]} = 0.0192$
- Intervalo de confianza al 95%: $0.08 \pm (1.96 \times 0.0192) = [0.042, 0.118]$

3.5.4 Distribución t de Student

Cuando la desviación estándar poblacional es desconocida y se estima mediante la desviación estándar muestral (s), se utiliza la distribución t:

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

Con grados de libertad: $gl = n - 1$

Características:

- Simétrica respecto al cero
- Mayor variabilidad que la normal estándar
- Converge a la normal cuando $n \rightarrow \infty$
- Colas más pesadas que la normal

3.5.5 Distribución Chi-cuadrado (χ^2)

Utilizada para inferencias sobre la varianza poblacional:

$$\chi^2 = \frac{(n - 1)s^2}{\sigma^2}$$

Con grados de libertad: $gl = n - 1$

Aplicación en control de calidad:

En la producción de alimento balanceado, se requiere que la variabilidad en el contenido proteico no exceda $\sigma^2 = 4$ ($\%$)². Una muestra de 25 lotes muestra $s^2 = 5.2$ ($\%$)².

$$\chi^2 = \frac{24 * 5.2}{4} = 31.2$$

Con $gl = 24$ y $\alpha = 0.05$, $\chi^2_{\text{crítico}} = 36.42$. Como $31.2 < 36.42$, no hay evidencia de variabilidad excesiva.

3.6 Teorema del límite central

3.6.1 Enunciado y significado

El Teorema del Límite Central (TLC) establece que, independientemente de la distribución de la población original, la distribución muestral de la media se aproxima a una distribución normal cuando el tamaño de muestra es suficientemente grande (Reynolds, 2023).

Formalmente, si $X_1, X_2, \dots, X_n$ son variables aleatorias independientes e idénticamente distribuidas con media μ y varianza σ^2 finita, entonces:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1) \text{ cuando } n \rightarrow \infty$$

3.6.2 Condiciones de aplicación

Tamaño de muestra mínimo:

- Para poblaciones simétricas: $n \geq 15$
- Para poblaciones moderadamente asimétricas: $n \geq 30$
- Para poblaciones muy asimétricas: $n \geq 40-50$

3.6.3 Implicaciones prácticas en veterinaria

1. **Validación de métodos paramétricos:** Permite aplicar pruebas t y ANOVA incluso cuando la población no es normal, siempre que n sea adecuado.
2. **Estimación de intervalos de confianza:** Facilita la construcción de intervalos para medias sin conocer la distribución poblacional.
3. **Diseño de muestreo:** Justifica el uso de fórmulas de tamaño de muestra basadas en la distribución normal.

3.6.4 Demostración práctica

Caso: Tiempo de recuperación postquirúrgica

El tiempo de recuperación tras ovariectomía en perras sigue una distribución exponencial con media $\mu = 3$ días y $\sigma = 3$ días (distribución muy asimétrica).

Para muestras de tamaño $n = 36$:

- Media de la distribución muestral: $\mu_{\bar{X}} = 3$ días
- Error estándar: $\sigma_{\bar{X}} = 3/\sqrt{36} = 0.5$ días

Por el TLC, $\bar{X} \sim N(3, 0.5^2)$

Probabilidad de que el tiempo promedio de recuperación en 36 perras sea menor a 2.8 días:

$$Z = \frac{2.8 - 3}{0.5} = -0.4$$

$$P(X^- < 2.8) = P(Z < -0.4) = 0.3446$$

3.6.5 Aplicaciones avanzadas

Comparación de medias independientes:

Cuando se comparan dos grupos con tamaños n_1 y n_2 suficientemente grandes, la diferencia de medias sigue aproximadamente:

$$(X_1 - X_2) \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

Ejemplo - Ganancia de peso en cerdos:

- Comparando dos dietas en cerdos de engorde:
- Dieta A: $n_1 = 40$, $\bar{X}_1 = 0.85$ kg/día, $s_1 = 0.12$ kg/día
- Dieta B: $n_2 = 45$, $\bar{X}_2 = 0.78$ kg/día, $s_2 = 0.15$ kg/día

Error estándar de la diferencia:

$$SE = \sqrt{\frac{0.12^2}{40} + \frac{0.15^2}{45}} = 0.0297$$

Estadístico Z:

$$Z = \frac{(0.85 - 0.78) - 0}{0.0297} = 2.36$$

Con $\alpha = 0.05$ bilateral, $Z_{\text{crítico}} = \pm 1.96$. Como $|2.36| > 1.96$, existe diferencia significativa entre las dietas.

3.7 Problemas propuestos

Nivel básico

1. En una granja avícola, la mortalidad diaria de pollos es del 2%. Si hay 500 pollos, calcule la probabilidad de que mueran exactamente 8 pollos en un día usando la distribución binomial.
2. La concentración de cortisol en suero de perros sanos sigue $N(4.5 \mu\text{g/dL}, 1.2 \mu\text{g/dL})$. ¿Qué porcentaje de perros sanos tendrá valores entre 3 y 6 $\mu\text{g/dL}$?
3. Una clínica veterinaria atiende en promedio 4 urgencias nocturnas por semana. Calcule la probabilidad de atender más de 6 urgencias en una semana usando la distribución de Poisson.

Nivel intermedio

1. Un test rápido para parvovirus canino tiene sensibilidad del 90% y especificidad del 95%. Si la prevalencia en cachorros no vacunados es del 30%, calcule e interprete los valores predictivos.
2. El tiempo de gestación en yeguas sigue $N(340 \text{ días}, 12 \text{ días})$. Si se toma una muestra de 25 yeguas: a) ¿Cuál es la probabilidad de que el promedio de gestación sea mayor a 345 días? b) ¿Entre qué valores se encuentra el 95% central de las medias muestrales?
3. La eficacia de una vacuna contra Newcastle en aves es del 88%. En un lote de 50 aves vacunadas: a) Calcule la probabilidad de que al menos 45 desarrollen inmunidad b) Determine el número esperado de aves inmunizadas y su desviación estándar

Nivel avanzado

1. En un estudio de mastitis bovina, se evalúan dos tratamientos. El tratamiento A tiene éxito en 72 de 90 vacas, mientras que el B en 65 de 85 vacas. Utilizando el TLC, determine si existe diferencia significativa entre las proporciones de éxito ($\alpha = 0.05$).
2. La producción de huevos en gallinas ponedoras tiene distribución asimétrica con $\mu = 280$ huevos/año y $\sigma = 45$ huevos/año. Si se seleccionan 49 gallinas al azar: a) Justifique el uso del TLC b) Calcule $P(270 < \bar{X} < 290)$ c) Determine el percentil 90 de la distribución muestral
3. En un programa de mejoramiento genético porcino, el peso a los 180 días sigue $N(105 \text{ kg}, 12 \text{ kg})$. Se seleccionan para reproducción los animales con peso superior a 115 kg. Si además se sabe que el 60% de los seleccionados son hembras: a) ¿Qué proporción de la población total es seleccionada? b) Si se evalúan 200 cerdos, ¿cuál es la probabilidad de obtener entre 30 y 45 animales aptos? c) De 50 animales seleccionados, ¿cuál es la probabilidad de obtener exactamente 30 hembras?

3.8 Actividades prácticas con software estadístico

```
# Distribución Normal - Valores de referencia hematológicos
# Hemoglobina en perros: N(15.5, 1.8) g/dL

# Visualización
library(ggplot2)
x <- seq(10, 21, 0.1)
y <- dnorm(x, mean = 15.5, sd = 1.8)
df <- data.frame(x, y)

ggplot(df, aes(x, y)) +
  geom_line(color = "blue", size = 1.2) +
  geom_area(data = subset(df, x >= 11.9 & x <= 19.1),
    aes(x, y), fill = "lightblue", alpha = 0.5) +
  labs(title = "Distribución de Hemoglobina en Perros Sanos",
    x = "Hemoglobina (g/dL)", y = "Densidad") +
  theme_minimal()

# Cálculo de probabilidades
prob_anemia <- pnorm(11.9, mean = 15.5, sd = 1.8)
print(paste("Probabilidad de anemia:", round(prob_anemia, 4)))

# Simulación del TLC
set.seed(123)
n_sim <- 1000
n_muestra <- 30
medias <- replicate(n_sim, mean(rexp(n_muestra, rate = 1/3)))
hist(medias, breaks = 30, main = "TLC: Distribución Exponencial →
Normal",
  xlab = "Media muestral", prob = TRUE)
curve(dnorm(x, mean = 3, sd = 3/sqrt(n_muestra)), add = TRUE, col =
"red", lwd = 2)
```

Figura 32. Práctica en R

3.9 Conclusiones del capítulo

La comprensión de los conceptos de probabilidad y las distribuciones estadísticas constituye la base fundamental para la aplicación correcta

de métodos estadísticos en medicina veterinaria y zootecnia. Los principios probabilísticos permiten al profesional veterinario cuantificar la incertidumbre inherente a los procesos biológicos, mientras que las distribuciones de probabilidad proporcionan modelos matemáticos para describir la variabilidad observada en poblaciones animales.

La distribución normal emerge como el modelo más importante en bioestadística veterinaria, no solo por su frecuente aparición en variables biológicas, sino también por su papel central en el Teorema del Límite Central. Este teorema justifica el uso de métodos paramétricos en una amplia gama de situaciones prácticas, incluso cuando las poblaciones originales no siguen distribuciones normales.

Las distribuciones binomiales y de Poisson resultan especialmente útiles en estudios epidemiológicos y evaluación de tratamientos, permitiendo modelar eventos discretos como prevalencia de enfermedades, eficacia vacunal y ocurrencia de casos. La correcta selección entre estas distribuciones depende de las características específicas del fenómeno bajo estudio.

El concepto de distribución muestral y el error estándar son fundamentales para comprender la precisión de las estimaciones y el diseño apropiado de estudios. El Teorema del Límite Central proporciona la justificación teórica para muchos procedimientos inferenciales utilizados rutinariamente en investigación veterinaria.

La aplicación práctica de estos conceptos requiere no solo comprensión teórica, sino también habilidad en el uso de software estadístico y capacidad para interpretar resultados en el contexto clínico o productivo

específico. Los ejemplos y ejercicios presentados en este capítulo buscan desarrollar estas competencias esenciales para el futuro profesional veterinario.

CAPÍTULO IV

4 INFERENCIA ESTADÍSTICA

4.1 Objetivos del capítulo

Al finalizar este capítulo, el estudiante será capaz de:

1. Comprender y aplicar los conceptos fundamentales de inferencia estadística en contextos veterinarios y zootécnicos
2. Calcular e interpretar estimaciones puntuales e intervalos de confianza para parámetros poblacionales
3. Diseñar y ejecutar pruebas de hipótesis apropiadas para investigación veterinaria
4. Evaluar y minimizar los errores tipo I y tipo II en estudios clínicos
5. Determinar la potencia estadística y calcular tamaños de muestra óptimos
6. Implementar análisis estadísticos usando R con datos veterinarios reales

4.2 Introducción

La inferencia estadística constituye el puente fundamental entre los datos observados en muestras y las conclusiones sobre poblaciones completas en medicina veterinaria. En la práctica clínica y la investigación zootécnica, raramente es posible examinar a todos los individuos de una población. Por ejemplo, evaluar la prevalencia de mastitis en todo el ganado lechero de una región, o determinar la eficacia de un nuevo tratamiento en todos los perros con displasia de cadera,

requeriría recursos prohibitivos. La inferencia estadística proporciona las herramientas metodológicas para extraer conclusiones válidas y confiables a partir de muestras representativas, cuantificando la incertidumbre inherente a este proceso.

Este capítulo desarrolla los conceptos fundamentales de la inferencia estadística con un enfoque aplicado a las ciencias veterinarias. A través de ejemplos prácticos con datos reales de estudios publicados en revistas indexadas, el estudiante aprenderá a implementar estas técnicas usando el software R, preparándose para realizar investigación de calidad y tomar decisiones clínicas basadas en evidencia.

4.3 Estimación puntual y por intervalos

4.3.1 Conceptos fundamentales

La estimación estadística es el proceso mediante el cual utilizamos información de una muestra para inferir características de una población. En medicina veterinaria, este proceso es esencial para determinar parámetros como la prevalencia de enfermedades, la eficacia de tratamientos, o los valores normales de referencia para análisis clínicos (Cummings et al., 2022).

Parámetro poblacional: Es una característica numérica de la población completa que generalmente desconocemos. Se denota con letras griegas (μ para media, σ para desviación estándar, π para proporción).

Estadístico muestral: Es una característica numérica calculada a partir de los datos de la muestra. Se denota con letras latinas ($\bar{x}$ para media

muestral, s para desviación estándar muestral, p para proporción muestral).

Estimador: Es la regla o fórmula que especifica cómo calcular una estimación del parámetro a partir de los datos muestrales.

4.3.2 Propiedades de los buenos estimadores

Un estimador debe cumplir ciertas propiedades deseables para ser considerado adecuado:

Insesgamiento: Un estimador es insesgado si su valor esperado es igual al parámetro que estima. Matemáticamente, $E(\theta) = \theta$, donde θ es el estimador y θ el parámetro (Candia B & Caiozzi A., 2005).

Eficiencia: Entre todos los estimadores insesgados, el más eficiente es aquel con menor varianza. Esto significa que sus estimaciones están más concentradas alrededor del valor verdadero.

Consistencia: Un estimador es consistente si converge al valor verdadero del parámetro cuando el tamaño de muestra aumenta hacia infinito.

Suficiencia: Un estimador es suficiente si utiliza toda la información relevante contenida en la muestra sobre el parámetro.

4.3.3 Estimación puntual

La estimación puntual proporciona un único valor como mejor aproximación del parámetro poblacional. Los estimadores puntuales más comunes en veterinaria incluyen:

Media muestral ($\bar{x}$): Estimador de la media poblacional μ

$$\bar{x} = \sum x_i / n$$

Varianza muestral (s^2): Estimador de la varianza poblacional σ^2

$$s^2 = \sum (x_i - \bar{x})^2 / (n-1)$$

Proporción muestral (p): Estimador de la proporción poblacional π

$$p = x/n$$

donde x es el número de éxitos en n observaciones.

4.3.4 Ejemplo práctico: Estimación del peso promedio en terneros

Un estudio evaluó el peso al destete de terneros Holstein en una explotación lechera. Se seleccionaron aleatoriamente 45 terneros y se registraron sus pesos en kilogramos (Cuestas et al., 2024).

```

# Datos de peso al destete (kg)
pesos <- c(185.2, 192.5, 178.8, 188.3, 195.1, 182.4, 190.7, 186.9,
          191.3, 184.6, 189.2, 187.5, 193.8, 180.9, 188.7, 194.2,
          183.1, 191.6, 186.3, 189.9, 185.8, 192.1, 187.2, 190.4,
          188.1, 184.3, 193.5, 181.7, 189.6, 186.8, 191.9, 183.9,
          188.5, 195.7, 182.2, 190.1, 187.8, 192.7, 185.4, 189.1,
          194.5, 181.3, 186.5, 191.1, 188.9)

# Estimación puntual
n <- length(pesos)
media_muestral <- mean(pesos)
varianza_muestral <- var(pesos)
desv_muestral <- sd(pesos)
cat("Tamaño de muestra: n =", n, "\n")
cat("Media muestral:  $\bar{x}$  =", round(media_muestral, 2), "kg\n")
cat("Varianza muestral:  $s^2$  =", round(varianza_muestral, 2), "kg2\n")
cat("Desviación estándar muestral: s =", round(desv_muestral, 2), "kg\n")

# Error estándar de la media
error_estandar <- desv_muestral/sqrt(n)
cat("Error estándar: SE =", round(error_estandar, 2), "kg\n")

```

Figura 33. Estimación del peso promedio en terneros.

4.3.5 Distribución muestral del estimador

Un concepto fundamental en inferencia es que los estimadores son variables aleatorias que tienen su propia distribución de probabilidad, llamada distribución muestral. Esta distribución describe cómo variarían las estimaciones si repitiéramos el muestreo infinitas veces (Calvache & Noguera, 2007).

Teorema del límite central: Para muestras suficientemente grandes ($n \geq 30$), la distribución muestral de la media se aproxima a una distribución normal, independientemente de la distribución de la población original:

$$\bar{x} \sim N(\mu, \sigma^2/n)$$

Este teorema es fundamental porque permite hacer inferencias incluso cuando desconocemos la distribución exacta de la población.

4.3.6 Estimación por intervalos

Mientras que la estimación puntual proporciona un único valor, la estimación por intervalos proporciona un rango de valores plausibles para el parámetro, cuantificando la incertidumbre de la estimación.

Un intervalo de estimación tiene la forma general:

[Estimador puntual - Margen de error, Estimador puntual + Margen de error]

La ventaja principal de los intervalos sobre las estimaciones puntuales es que proporcionan información sobre la precisión de la estimación. Un intervalo estrecho indica alta precisión, mientras que un intervalo amplio sugiere mayor incertidumbre.

4.3.7 Ejemplo aplicado: Prevalencia de Brucelosis Bovina

Un estudio epidemiológico investigó la prevalencia de brucelosis en ganado bovino de una región ganadera. Se analizaron 380 animales mediante la prueba de Rosa de Bengala.

```

# Datos del estudio
n_animales <- 380
positivos <- 42
# Estimación puntual de prevalencia
prevalencia <- positivos/n_animales
cat("Prevalencia estimada: p =", round(prevalencia*100, 2), "%\n")
# Error estándar para proporción
error_est_prop <- sqrt((prevalencia * (1-prevalencia))/n_animales)
cat("Error estándar: SE =", round(error_est_prop*100, 3), "%\n")
# Visualización de la distribución muestral
library(ggplot2)
set.seed(123)
# Simulación de distribución muestral
simulaciones <- rbinom(10000, n_animales, prevalencia)/n_animales
datos_sim <- data.frame(prevalencia_simulada = simulaciones)

ggplot(datos_sim, aes(x = prevalencia_simulada)) +
  geom_histogram(aes(y = ..density..), bins = 30,
    fill = "lightblue", color = "black", alpha = 0.7) +
  geom_vline(xintercept = prevalencia, color = "red",
    linetype = "dashed", size = 1) +
  stat_function(fun = dnorm,
    args = list(mean = prevalencia, sd = error_est_prop),
    color = "blue", size = 1) +
  labs(title = "Distribución Muestral de la Prevalencia de Brucelosis",
    x = "Prevalencia estimada",
    y = "Densidad") +
  theme_minimal()

```

Figura 34. Prevalencia de Brucelosis Bovina.

4.4 Intervalos de confianza

4.4.1 Concepto e interpretación

Un intervalo de confianza (IC) es un rango de valores que, con un nivel de confianza especificado, contiene el verdadero valor del parámetro poblacional. El nivel de confianza más común es 95%, aunque en veterinaria también se utilizan niveles de 90% y 99% según el contexto (Blas Vilca, 2022).

Interpretación correcta: Si construyéramos 100 intervalos de confianza del 95% a partir de 100 muestras diferentes, aproximadamente 95 de ellos contendrían el verdadero valor del parámetro.

Interpretación incorrecta común: NO es correcto decir que existe un 95% de probabilidad de que el parámetro esté en el intervalo calculado. El parámetro es fijo; el intervalo es el que varía con cada muestra.

4.4.2 Construcción de intervalos de confianza

La estructura general de un intervalo de confianza es:

$$IC = \text{Estimador puntual} \pm (\text{Valor crítico} \times \text{Error estándar})$$

Intervalo de confianza para la media con σ conocida:

$$IC(\mu) = \bar{x} \pm z(\alpha/2) \times (\sigma/\sqrt{n})$$

Intervalo de confianza para la media con σ desconocida:

$$IC(\mu) = \bar{x} \pm t(\alpha/2, n-1) \times (s/\sqrt{n})$$

Intervalo de confianza para una proporción:

$$IC(\pi) = p \pm z(\alpha/2) \times \sqrt{[p(1-p)/n]}$$

4.4.3 Ejemplo práctico: Concentración de cortisol en perros con estrés

Un estudio evaluó los niveles de cortisol sérico en perros sometidos a hospitalización. Se analizaron muestras de 28 perros durante las primeras 24 horas de internación.

```

# Datos de cortisol sérico (µg/dL)
cortisol <- c(8.2, 9.5, 7.8, 10.3, 8.9, 9.1, 7.5, 8.7, 9.8, 8.3,
             9.2, 7.9, 8.6, 10.1, 8.4, 9.6, 8.1, 9.3, 7.7, 8.8,
             9.4, 8.5, 10.2, 7.6, 9.0, 8.9, 9.7, 8.2)

# Estadísticas descriptivas
n <- length(cortisol)
media <- mean(cortisol)
desv_est <- sd(cortisol)
error_est <- desv_est/sqrt(n)
# Intervalo de confianza del 95% para la media
alpha <- 0.05
t_critico <- qt(1 - alpha/2, df = n - 1)
IC_inferior <- media - t_critico * error_est
IC_superior <- media + t_critico * error_est

cat("Media muestral:", round(media, 2), "µg/dL\n")
cat("Intervalo de confianza del 95%: [",
    round(IC_inferior, 2), ", ", round(IC_superior, 2), "] µg/dL\n")

# Función en R para IC
t.test(cortisol, conf.level = 0.95)

# Visualización del IC
library(ggplot2)
datos_cortisol <- data.frame(
  grupo = "Cortisol",
  media = media,
  ic_inf = IC_inferior,
  ic_sup = IC_superior
)
ggplot(datos_cortisol, aes(x = grupo, y = media)) +
  geom_point(size = 3) +
  geom_errorbar(aes(ymin = ic_inf, ymax = ic_sup),
    width = 0.1, size = 1) +
  geom_hline(yintercept = 6, linetype = "dashed",
    color = "red", alpha = 0.5) +
  annotate("text", x = 0.7, y = 6.2,
    label = "Valor normal máximo", color = "red") +
  labs(title = "Nivel de Cortisol en Perros Hospitalizados",
    subtitle = "Media e Intervalo de Confianza del 95%",
    y = "Cortisol sérico (µg/dL)",
    x = "") +
  theme_minimal() +
  ylim(5, 11)

```

Figura 35. Concentración de cortisol en perros con estrés

4.4.4 Intervalos de confianza para diferencias

En investigación veterinaria, frecuentemente necesitamos comparar dos grupos. Los intervalos de confianza para diferencias son especialmente útiles porque permiten evaluar tanto la significancia estadística como la relevancia clínica (Castañeda & Fabián Gil, 2004).

IC para diferencia de medias (muestras independientes):

$$IC(\mu_1 - \mu_2) = (\bar{x}_1 - \bar{x}_2) \pm t(\alpha/2, gl) \times SE(\bar{x}_1 - \bar{x}_2)$$

IC para diferencia de proporciones:

$$IC(\pi_1 - \pi_2) = (p_1 - p_2) \pm z(\alpha/2) \times SE(p_1 - p_2)$$

4.4.5 Caso clínico: Comparación de tratamientos para mastitis bovina

Un ensayo clínico comparó dos tratamientos antibióticos para mastitis clínica en vacas lecheras. Se evaluó la tasa de curación bacteriológica a los 21 días post-tratamiento.

```

# Datos del ensayo clínico
tratamiento_A <- c(n = 85, curacion = 68)
tratamiento_B <- c(n = 92, curacion = 71)
# Proporciones de curación
p_A <- tratamiento_A["curacion"]/tratamiento_A["n"]
p_B <- tratamiento_B["curacion"]/tratamiento_B["n"]
# Diferencia de proporciones
diferencia <- p_A - p_B
# Error estándar de la diferencia
SE_dif <- sqrt((p_A*(1-p_A)/tratamiento_A["n"]) +
               (p_B*(1-p_B)/tratamiento_B["n"]))
# IC del 95% para la diferencia
z_critico <- qnorm(0.975)
IC_dif_inf <- diferencia - z_critico * SE_dif
IC_dif_sup <- diferencia + z_critico * SE_dif
cat("Tasa de curación Tratamiento A:", round(p_A*100, 1), "%\n")
cat("Tasa de curación Tratamiento B:", round(p_B*100, 1), "%\n")
cat("Diferencia:", round(diferencia*100, 1), "%\n")
cat("IC 95% para la diferencia: [", round(IC_dif_inf*100, 1),
    "%,", round(IC_dif_sup*100, 1), "%]\n")
# Prueba usando prop.test
resultado <- prop.test(c(tratamiento_A["curacion"], tratamiento_B["curacion"]),
                       c(tratamiento_A["n"], tratamiento_B["n"]),
                       conf.level = 0.95)
print(resultado)
# Visualización comparativa
library(ggplot2)
datos_trat <- data.frame(
  Tratamiento = c("A", "B"),
  Tasa_curacion = c(p_A, p_B) * 100,
  IC_inf = c(p_A - 1.96*sqrt(p_A*(1-p_A)/tratamiento_A["n"]),
             p_B - 1.96*sqrt(p_B*(1-p_B)/tratamiento_B["n"])) * 100,
  IC_sup = c(p_A + 1.96*sqrt(p_A*(1-p_A)/tratamiento_A["n"]),
             p_B + 1.96*sqrt(p_B*(1-p_B)/tratamiento_B["n"])) * 100
)
ggplot(datos_trat, aes(x = Tratamiento, y = Tasa_curacion,
                       color = Tratamiento)) +
  geom_point(size = 4) + geom_errorbar(aes(ymin = IC_inf, ymax = IC_sup),
                                       width = 0.1, size = 1) +
  geom_hline(yintercept = 75, linetype = "dashed",
             alpha = 0.5) +
  labs(title = "Eficacia de Tratamientos para Mastitis Bovina",
        subtitle = "Tasas de curación bacteriológica con IC 95%",
        y = "Tasa de curación (%)",
        x = "Tratamiento") + theme_minimal() + scale_color_manual(values = c("blue",
"red")) + theme(legend.position = "none")

```

Figura 36. Comparación de tratamientos para mastitis bovina.

4.4.6 Factores que afectan la amplitud del IC

La precisión de un intervalo de confianza, reflejada en su amplitud, depende de varios factores:

1. **Tamaño de muestra (n):** Mayor n produce intervalos más estrechos
2. **Variabilidad (σ o s):** Mayor variabilidad produce intervalos más amplios
3. **Nivel de confianza:** Mayor confianza (99% vs 95%) produce intervalos más amplios

```

# Demostración del efecto del tamaño de muestra
set.seed(123)
media_pob <- 100
desv_pob <- 15

# Función para calcular IC
calcular_IC <- function(n, nivel_conf = 0.95) {
  muestra <- rnorm(n, media_pob, desv_pob)
  media <- mean(muestra)
  error_est <- sd(muestra)/sqrt(n)
  t_crit <- qt((1 + nivel_conf)/2, df = n - 1)
  margen <- t_crit * error_est
  return(c(inferior = media - margen,
           superior = media + margen,
           amplitud = 2 * margen))
}

# Diferentes tamaños de muestra
tamaños <- c(10, 25, 50, 100, 200)
resultados <- data.frame()

for(n in tamaños) {
  ic <- calcular_IC(n)
  resultados <- rbind(resultados,
                     data.frame(n = n,
                               inferior = ic[1],
                               superior = ic[2],
                               amplitud = ic[3]))
}

# Visualización
library(ggplot2)
ggplot(resultados, aes(x = factor(n), y = 100)) +
  geom_errorbar(aes(ymin = inferior, ymax = superior),
               width = 0.2, size = 1) +
  geom_point(size = 3) +
  geom_hline(yintercept = 100, linetype = "dashed",
            color = "red", alpha = 0.5) +
  labs(title = "Efecto del Tamaño de Muestra en la Precisión del IC",
       x = "Tamaño de muestra (n)",
       y = "Peso promedio (kg)") +
  theme_minimal()

```

Figura 37. Factores que afectan la amplitud del IC.

4.5 Pruebas de hipótesis

4.5.1 Fundamentos conceptuales

Una prueba de hipótesis es un procedimiento estadístico para evaluar afirmaciones sobre parámetros poblacionales basándose en evidencia muestral (Peña & Pérez, 2015). En medicina veterinaria, las pruebas de hipótesis permiten responder preguntas como:

- ¿Es efectivo este nuevo tratamiento?
- ¿Difiere la prevalencia de una enfermedad entre dos regiones?
- ¿Existe asociación entre un factor de riesgo y una patología?

El proceso sigue una lógica de refutación: comenzamos asumiendo que no hay efecto (hipótesis nula) y evaluamos si los datos proporcionan suficiente evidencia para rechazar esta suposición.

4.5.2 Elementos de una prueba de hipótesis

Hipótesis nula (H_0): Afirmación de no efecto o no diferencia. Representa el statu quo.

Hipótesis alternativa (H_1 o H_a): Afirmación que contradice H_0 . Representa lo que queremos demostrar.

Estadístico de prueba: Medida calculada de los datos que cuantifica la discrepancia entre lo observado y lo esperado bajo H_0 .

Valor p: Probabilidad de obtener un resultado tan extremo o más que el observado, asumiendo que H_0 es verdadera.

Nivel de significancia (α): Probabilidad máxima de rechazar H_0 cuando es verdadera. Típicamente $\alpha = 0.05$.

Regla de decisión: Si $p < \alpha$, rechazamos H_0 ; si $p \geq \alpha$, no rechazamos H_0 .

4.5.3 Tipos de pruebas de hipótesis

Las pruebas pueden clasificarse según la dirección de la hipótesis alternativa:

Prueba bilateral (dos colas):

- $H_0: \mu = \mu_0$
- $H_1: \mu \neq \mu_0$
- Usada cuando no hay expectativa previa sobre la dirección del efecto

Prueba unilateral derecha:

- $H_0: \mu \leq \mu_0$
- $H_1: \mu > \mu_0$
- Usada cuando esperamos que el parámetro sea mayor

Prueba unilateral izquierda:

- $H_0: \mu \geq \mu_0$
- $H_1: \mu < \mu_0$
- Usada cuando esperamos que el parámetro sea menor

4.5.4 Ejemplo práctico: Evaluación de un suplemento nutricional

Un productor avícola desea evaluar si un nuevo suplemento nutricional aumenta el peso de pollos de engorde a los 42 días. El peso promedio histórico es de 2.5 kg.

```

# Datos de peso (kg) con el nuevo suplemento
pesos_pollos <- c(2.68, 2.55, 2.73, 2.61, 2.59, 2.71, 2.64, 2.67,
  2.58, 2.70, 2.63, 2.72, 2.56, 2.69, 2.65, 2.74,
  2.60, 2.66, 2.57, 2.71, 2.62, 2.68, 2.75, 2.59,
  2.70, 2.64, 2.72, 2.61, 2.67, 2.58)

# Hipótesis
# H0:  $\mu = 2.5$  kg (el suplemento no tiene efecto)
# H1:  $\mu > 2.5$  kg (el suplemento aumenta el peso)
mu_0 <- 2.5
n <- length(pesos_pollos)
media_muestra <- mean(pesos_pollos)
desv_muestra <- sd(pesos_pollos)
# Estadístico de prueba t
t_calculado <- (media_muestra - mu_0)/(desv_muestra/sqrt(n))
# Valor p para prueba unilateral derecha
valor_p <- pt(t_calculado, df = n - 1, lower.tail = FALSE)
# Resultados
cat("Estadísticas descriptivas:\n")
cat(" Media muestral:", round(media_muestra, 3), "kg\n")
cat(" Desviación estándar:", round(desv_muestra, 3), "kg\n")
cat("\nPrueba de hipótesis:\n")
cat(" Estadístico t:", round(t_calculado, 3), "\n")
cat(" Valor p:", round(valor_p, 4), "\n")
cat(" Decisión ( $\alpha = 0.05$ ):", ifelse(valor_p < 0.05, "Rechazar H0", "No rechazar H0"), "\n")
# Prueba t usando función de R
resultado_t <- t.test(pesos_pollos, mu = mu_0,
  alternative = "greater") print(resultado_t)
# Visualización de la prueba
library(ggplot2)
x_vals <- seq(-4, 5, length.out = 1000)
y_vals <- dt(x_vals, df = n - 1)
datos_dist <- data.frame(x = x_vals, y = y_vals)
datos_dist$region <- ifelse(datos_dist$x > t_calculado, "Rechazo", "No
rechazo") ggplot(datos_dist, aes(x = x, y = y)) +
  geom_line(size = 1) + geom_area(data = subset(datos_dist, region ==
"Rechazo"), aes(x = x, y = y), fill = "red", alpha = 0.3) +
  geom_vline(xintercept = t_calculado, color = "blue", linetype = "dashed",
size = 1) + geom_vline(xintercept = qt(0.95, df = n - 1), color = "red",
linetype = "solid", size = 1) + annotate("text", x = t_calculado + 0.3, y = 0.3,
label = paste("t =", round(t_calculado, 2)),
color = "blue") + annotate("text", x = 3, y = 0.05,
label = paste("p =", round(valor_p, 4)),
color = "red") + labs(title = "Prueba t Unilateral: Efecto del Suplemento
Nutricional", x = "Estadístico t", y = "Densidad") + theme_minimal()

```

Figura 38. Evaluación de un suplemento nutricional.

4.5.5 Pruebas para diferentes situaciones

Prueba t para dos muestras independientes

Compara las medias de dos grupos independientes. Común en ensayos clínicos veterinarios.

```
# Ejemplo: Comparación de ganancia de peso entre dos dietas
dieta_A <- c(325, 315, 340, 328, 335, 342, 318, 330, 337, 322,
            345, 320, 338, 326, 333)
dieta_B <- c(310, 305, 318, 312, 308, 315, 302, 320, 311, 306,
            319, 304, 316, 309, 313)

# Prueba t de dos muestras
t.test(dieta_A, dieta_B, var.equal = TRUE)

# Verificación de supuestos
shapiro.test(dieta_A) # Normalidad
shapiro.test(dieta_B) # Normalidad
var.test(dieta_A, dieta_B) # Homogeneidad de varianzas
```

Figura 39. Prueba t para dos muestras independientes.

Prueba t pareada

Para muestras relacionadas, como mediciones antes-después en los mismos animales.

```
# Ejemplo: Frecuencia cardíaca antes y después de sedación
FC_antes <- c(145, 138, 142, 150, 135, 148, 140, 143, 139, 147)
FC_despues <- c(95, 92, 98, 102, 88, 96, 91, 94, 90, 97)

# Prueba t pareada
t.test(FC_antes, FC_despues, paired = TRUE)

# Visualización del cambio
diferencias <- FC_antes - FC_despues
hist(diferencias, main = "Distribución de las Diferencias",
      xlab = "Reducción en FC (lpm)", col = "lightblue")
abline(v = mean(diferencias), col = "red", lwd = 2)
```

Figura 40. Prueba t pareada.

Prueba de proporciones

Para comparar proporciones o prevalencias entre grupos.

```
# Ejemplo: Efectividad de vacunas
vacuna_A <- c(protegidos = 142, total = 160)
vacuna_B <- c(protegidos = 128, total = 155)

# Prueba de proporciones
prop.test(c(vacuna_A["protegidos"], vacuna_B["protegidos"]),
          c(vacuna_A["total"], vacuna_B["total"]))
```

Figura 41. Prueba de proporciones.

4.5.6 Relación entre intervalos de confianza y pruebas de hipótesis

Existe una correspondencia directa entre intervalos de confianza y pruebas de hipótesis bilaterales:

- Si el IC del 95% no contiene el valor hipotético, rechazamos H_0 con $\alpha = 0.05$

- Si el IC del 95% contiene el valor hipotético, no rechazamos H_0

Esta relación hace que los IC sean más informativos que las pruebas de hipótesis solas, pues proporcionan el rango de valores plausibles además de la decisión estadística.

```
# Demostración de la relación IC - Prueba de hipótesis
set.seed(456)
muestra <- rnorm(30, mean = 52, sd = 8)

# Prueba de hipótesis:  $H_0: \mu = 50$ 
prueba <- t.test(muestra, mu = 50)

# Intervalo de confianza
ic <- prueba$conf.int

cat("Intervalo de confianza 95%: [", round(ic[1], 2),
    ", ", round(ic[2], 2), "]\n")
cat("¿El IC contiene a 50?:",
    ifelse(50 >= ic[1] & 50 <= ic[2], "Sí", "No"), "\n")
cat("Valor p de la prueba:", round(prueba$p.value, 4), "\n")
cat("Decisión:", ifelse(prueba$p.value < 0.05,
    "Rechazar  $H_0$ ", "No rechazar  $H_0$ "), "\n")
```

Figura 42. Relación entre intervalos de confianza y pruebas de hipótesis.

4.6 Error tipo I y tipo II

4.6.1 Definición y consecuencias

En toda prueba de hipótesis existe la posibilidad de cometer errores. La comprensión de estos errores es fundamental para el diseño e interpretación de estudios veterinarios.

Error tipo I (α): Rechazar H_0 cuando es verdadera (falso positivo)

- Consecuencia: Adoptar un tratamiento inefectivo

- Probabilidad controlada por el nivel de significancia α

Error tipo II (β): No rechazar H_0 cuando es falsa (falso negativo)

- Consecuencia: No detectar un tratamiento efectivo
- Probabilidad relacionada con la potencia estadística ($1 - \beta$)

Tabla 23. Tabla de decisiones y errores.

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 ($1-\beta$)

4.6.2 Factores que afectan los errores

Los errores tipo I y II están inversamente relacionados: reducir uno aumenta el otro. Los factores que influyen incluyen:

1. **Nivel de significancia (α):** Menor α reduce error tipo I pero aumenta error tipo II
2. **Tamaño de muestra (n):** Mayor n reduce ambos errores
3. **Variabilidad (σ):** Menor variabilidad reduce ambos errores
4. **Tamaño del efecto (δ):** Mayor efecto reduce error tipo II

4.6.3 Ejemplo aplicado: Detección de residuos antibióticos

Un laboratorio evalúa un nuevo test rápido para detectar residuos de antibióticos en leche. Se analizan 500 muestras con estado conocido.

```

# Resultados del test
verdaderos_positivos <- 85
falsos_positivos <- 12
verdaderos_negativos <- 388
falsos_negativos <- 15

# Cálculo de errores
total_positivos_reales <- verdaderos_positivos + falsos_negativos
total_negativos_reales <- verdaderos_negativos + falsos_positivos

error_tipo_I <- falsos_positivos / total_negativos_reales
error_tipo_II <- falsos_negativos / total_positivos_reales

sensibilidad <- verdaderos_positivos / total_positivos_reales
especificidad <- verdaderos_negativos / total_negativos_reales

cat("Error Tipo I ( $\alpha$ ):", round(error_tipo_I * 100, 2), "%\n")
cat("Error Tipo II ( $\beta$ ):", round(error_tipo_II * 100, 2), "%\n")
cat("Sensibilidad ( $1-\beta$ ):", round(sensibilidad * 100, 2), "%\n")
cat("Especificidad ( $1-\alpha$ ):", round(especificidad * 100, 2), "%\n")

# Matriz de confusión
library(ggplot2)
matriz <- matrix(c(verdaderos_positivos, falsos_positivos,
                   falsos_negativos, verdaderos_negativos),
                 nrow = 2, byrow = TRUE)
colnames(matriz) <- c("Test +", "Test -")
rownames(matriz) <- c("Real +", "Real -")

# Visualización
library(reshape2)
matriz_long <- melt(matriz)
colnames(matriz_long) <- c("Real", "Test", "Frecuencia")

ggplot(matriz_long, aes(x = Test, y = Real, fill = Frecuencia)) +
  geom_tile() +
  geom_text(aes(label = Frecuencia), size = 6) +
  scale_fill_gradient(low = "white", high = "darkblue") +
  labs(title = "Matriz de Confusión: Test de Residuos Antibióticos",
       x = "Resultado del Test",
       y = "Estado Real") +
  theme_minimal()

```

Figura 43. Detección de residuos antibióticos.

4.6.4 Control de errores en estudios veterinarios

En la práctica veterinaria, la elección del balance entre errores depende del contexto:

Situaciones donde minimizar Error Tipo I es prioritario:

- Aprobación de nuevos fármacos
- Declaración de zonas libres de enfermedad
- Certificación de productos para exportación

Situaciones donde minimizar Error Tipo II es prioritario:

- Screening de enfermedades graves
- Detección temprana de brotes epidémicos
- Identificación de animales portadores

```

# Simulación: Trade-off entre errores
set.seed(789)

# Parámetros
mu_0 <- 100 # Media bajo Ho
mu_1 <- 105 # Media real
sigma <- 10
n <- 30
alphas <- seq(0.01, 0.20, 0.01)

resultados <- data.frame()

for(alpha in alphas) {
  # Valor crítico
  z_crit <- qnorm(1 - alpha)

  # Error tipo II
  beta <- pnorm(z_crit, mean = (mu_1 - mu_0)/(sigma/sqrt(n)),
               sd = 1, lower.tail = TRUE)

  resultados <- rbind(resultados,
                     data.frame(alpha = alpha,
                               beta = beta,
                               potencia = 1 - beta))
}

# Visualización del trade-off
library(ggplot2)
ggplot(resultados, aes(x = alpha, y = beta)) +
  geom_line(size = 1.5, color = "red") +
  geom_point(size = 2) +
  labs(title = "Trade-off entre Error Tipo I y Error Tipo II",
       x = "Error Tipo I ( $\alpha$ )",
       y = "Error Tipo II ( $\beta$ )") +
  theme_minimal() +
  scale_x_continuous(labels = scales::percent) +
  scale_y_continuous(labels = scales::percent)

```

Figura 44. Control de errores en estudios veterinarios.

4.6.5 Corrección por comparaciones múltiples

Cuando realizamos múltiples pruebas simultáneamente, la probabilidad de cometer al menos un error tipo I aumenta. En estudios veterinarios con múltiples variables, esto es crítico.

Tasa de error familiar (FWER): Probabilidad de al menos un falso positivo.

$$\text{FWER} = 1 - (1 - \alpha)^m$$

donde m es el número de pruebas.

Corrección de Bonferroni: Método conservador que ajusta α

$$\alpha_{\text{ajustado}} = \alpha / m$$

```

# Ejemplo: Análisis de 10 parámetros sanguíneos
n_pruebas <- 10
alpha_nominal <- 0.05

# Sin corrección
FWER <- 1 - (1 - alpha_nominal)^n_pruebas
cat("Probabilidad de al menos un falso positivo sin corrección:",
    round(FWER * 100, 1), "%\n")

# Con corrección de Bonferroni
alpha_bonferroni <- alpha_nominal / n_pruebas
cat("Alpha ajustado (Bonferroni):", alpha_bonferroni, "\n")

# Simulación de p-valores
set.seed(321)
p_valores <- c(0.003, 0.012, 0.024, 0.041, 0.058,
               0.089, 0.134, 0.245, 0.456, 0.789)

# Decisiones
decisiones <- data.frame(
  Parametro = paste("Parámetro", 1:10),
  p_valor = p_valores,
  Sin_correccion = ifelse(p_valores < alpha_nominal, "Significativo", "No sig."),
  Bonferroni = ifelse(p_valores < alpha_bonferroni, "Significativo", "No sig."),
  Holm = ifelse(p.adjust(p_valores, method = "holm") < alpha_nominal,
                "Significativo", "No sig.")
)

print(decisiones)

```

Figura 45. Corrección por comparaciones múltiples.

4.7 Potencia estadística

4.7.1 Concepto e importancia

La potencia estadística es la probabilidad de rechazar correctamente una hipótesis nula falsa (Corcuera Valverde, 1994). Representa la capacidad de un estudio para detectar un efecto cuando realmente existe.

$$\text{Potencia} = 1 - \beta = P(\text{Rechazar } H_0 \mid H_0 \text{ es falsa})$$

Una potencia adecuada (típicamente $\geq 80\%$) es esencial para:

- Justificar éticamente el uso de animales en investigación
- Evitar desperdiciar recursos en estudios inconclusos
- Asegurar que efectos clínicamente importantes sean detectados

4.7.2 Factores que determinan la potencia

La potencia depende de cuatro factores interrelacionados:

1. **Tamaño del efecto (δ):** Diferencia entre H_0 y H_1
2. **Nivel de significancia (α):** Típicamente 0.05
3. **Tamaño de muestra (n):** Mayor n , mayor potencia
4. **Variabilidad (σ):** Menor variabilidad, mayor potencia

4.7.3 Cálculo de potencia para diferentes diseños

Potencia para prueba t de una muestra:

$$\text{Potencia} = 1 - \beta = P(\text{Rechazar } H_0 \mid H_0 \text{ es falsa})$$

Una potencia adecuada (típicamente $\geq 80\%$) es esencial para:

- Justificar éticamente el uso de animales en investigación
- Evitar desperdiciar recursos en estudios inconclusos
- Asegurar que efectos clínicamente importantes sean detectados

4.7.4 Factores que determinan la potencia

La potencia depende de cuatro factores interrelacionados:

1. **Tamaño del efecto (δ):** Diferencia entre H_0 y H_1
2. **Nivel de significancia (α):** Típicamente 0.05
3. **Tamaño de muestra (n):** Mayor n , mayor potencia
4. **Variabilidad (σ):** Menor variabilidad, mayor potencia

4.7.5 Cálculo de potencia para diferentes diseños

Potencia para prueba t de una muestra:

```

# Función para calcular potencia
library(pwr)

# Parámetros del estudio
tamaño_efecto <- 0.5 # d de Cohen
n <- 30
alpha <- 0.05

# Cálculo de potencia
potencia <- pwr.t.test(n = n,
                      d = tamaño_efecto,
                      sig.level = alpha,
                      type = "one.sample")

cat("Potencia del estudio:", round(potencia$power * 100, 1), "%\n")

# Curva de potencia
ns <- seq(10, 100, by = 5)
potencias <- sapply(ns, function(n) {
  pwr.t.test(n = n, d = tamaño_efecto,
            sig.level = alpha, type = "one.sample")$power
})

datos_potencia <- data.frame(n = ns, Potencia = potencias)

library(ggplot2)
ggplot(datos_potencia, aes(x = n, y = Potencia)) +
  geom_line(size = 1.5, color = "blue") +
  geom_hline(yintercept = 0.80, linetype = "dashed", color = "red") +
  annotate("text", x = 20, y = 0.82, label = "Potencia = 80%", color =
"red") +
  labs(title = "Curva de Potencia: Efecto del Tamaño de Muestra",
       x = "Tamaño de muestra (n)",
       y = "Potencia estadística") +
  theme_minimal() +
  scale_y_continuous(labels = scales::percent)

```

Figura 46. Potencia para prueba t de una muestra.

4.7.6 Ejemplo práctico: Ensayo de eficacia vacunal

Un investigador planea un estudio para evaluar una nueva vacuna contra leptospirosis canina. Basándose en estudios previos, espera reducir la incidencia del 25% al 10%.

```

# Parámetros del estudio
p1 <- 0.25 # Incidencia sin vacuna
p2 <- 0.10 # Incidencia esperada con vacuna
alpha <- 0.05
potencia_deseada <- 0.80

# Cálculo del tamaño del efecto (h de Cohen)
h <- 2 * asin(sqrt(p1)) - 2 * asin(sqrt(p2))

# Tamaño de muestra necesario
library(pwr)
resultado <- pwr.2p.test(h = h,
  sig.level = alpha,
  power = potencia_deseada)

n_por_grupo <- ceiling(resultado$n)
cat("Tamaño del efecto (h):", round(h, 3), "\n")
cat("Animales necesarios por grupo:", n_por_grupo, "\n")
cat("Total de animales:", 2 * n_por_grupo, "\n")

# Análisis de sensibilidad: potencia para diferentes escenarios
escenarios <- data.frame(
  p2 = seq(0.05, 0.20, 0.01)
)

escenarios$h <- 2 * asin(sqrt(p1)) - 2 * asin(sqrt(escenarios$p2))
escenarios$n <- sapply(escenarios$h, function(h) {
  ceiling(pwr.2p.test(h = h, sig.level = alpha,
    power = potencia_deseada)$n)
})

# Visualización
ggplot(escenarios, aes(x = p2 * 100, y = n)) +
  geom_line(size = 1.5) +
  geom_point(data = data.frame(p2 = 10, n = n_por_grupo),
    aes(x = p2, y = n), color = "red", size = 3) +
  labs(title = "Tamaño de Muestra según Eficacia Esperada de la Vacuna",
    x = "Incidencia esperada con vacuna (%)",
    y = "n por grupo") +
  theme_minimal()

```

Figura 47. Ensayo de eficacia vacunal.

4.7.7 *Análisis de potencia Post-Hoc*

Aunque controversial, el análisis de potencia post-hoc puede ser útil para interpretar resultados negativos y planificar estudios futuros.

```
# Ejemplo: Estudio con resultado no significativo
grupo_control <- c(45.2, 43.8, 46.1, 44.5, 45.9, 43.2, 46.8, 44.1,
                  45.5, 43.9, 46.3, 44.7)
grupo_tratado <- c(46.1, 44.9, 47.2, 45.3, 46.8, 44.5, 47.5, 45.1,
                  46.4, 45.0, 47.1, 45.7)

# Prueba t
resultado <- t.test(grupo_tratado, grupo_control, var.equal = TRUE)
cat("Valor p:", round(resultado$p.value, 3), "\n")

# Cálculo del tamaño del efecto observado
diferencia_obs <- mean(grupo_tratado) - mean(grupo_control)
pool_sd <- sqrt(((length(grupo_control)-1)*var(grupo_control) +
               (length(grupo_tratado)-1)*var(grupo_tratado)) /
               (length(grupo_control) + length(grupo_tratado) - 2))
d_observado <- diferencia_obs / pool_sd

# Potencia post-hoc
potencia_post <- pwr.t.test(n = length(grupo_control),
                           d = d_observado,
                           sig.level = 0.05,
                           type = "two.sample")$power

cat("Tamaño del efecto observado (d):", round(d_observado, 3), "\n")
cat("Potencia para detectar este efecto:", round(potencia_post * 100, 1), "%\n")

# ¿Cuántos animales necesitaríamos para 80% de potencia?
n_necesario <- pwr.t.test(d = d_observado,
                          sig.level = 0.05,
                          power = 0.80,
                          type = "two.sample")$n

cat("n necesario para 80% potencia:", ceiling(n_necesario), "por grupo\n")
```

Figura 48. Análisis de potencia Post-Hoc.

4.7.8 Potencia en diseños complejos

En estudios veterinarios con diseños jerárquicos o medidas repetidas, el cálculo de potencia requiere consideraciones adicionales (Bologna, 2022).

```
# Ejemplo: Diseño con estructura jerárquica (animales en granjas)
library(lme4)
library(simr)

# Simulación de datos con estructura jerárquica
set.seed(654)
n_granjas <- 10
n_animales_granja <- 15
efecto_tratamiento <- 2.5
sigma_granja <- 3
sigma_residual <- 5

# Generar datos simulados
datos_sim <- expand.grid(
  animal = 1:n_animales_granja,
  granja = factor(1:n_granjas)
)
datos_sim$tratamiento <- rep(c("Control", "Tratado"),
                             each = n_granjas * n_animales_granja / 2)

# Modelo mixto
modelo_sim <- makeLmer(y ~ tratamiento + (1|granja),
                      fixef = c(50, efecto_tratamiento),
                      VarCorr = sigma_granja,
                      sigma = sigma_residual,
                      data = datos_sim)

# Análisis de potencia
potencia_sim <- powerSim(modelo_sim, nsim = 100, progress = FALSE)
print(potencia_sim)
```

Figura 49. Potencia en diseños complejos.

4.8 Tamaño de muestra

4.8.1 Importancia del cálculo a priori

El cálculo del tamaño de muestra antes de iniciar un estudio es fundamental por razones:

Éticas: Cumplimiento del principio de las 3Rs (Reducir número de animales)

Científicas: Asegurar potencia adecuada para detectar efectos importantes

Económicas: Optimizar recursos evitando estudios subdimensionados o sobredimensionados

Regulatorias: Requerido por comités de ética y agencias reguladoras

4.8.2 Fórmulas básicas para tamaño de muestra

Para estimar una media:

$$n = (z_{\alpha/2} \times \sigma / E)^2$$

donde E es el margen de error tolerado.

Para estimar una proporción:

$$n = z_{\alpha/2}^2 \times p(1-p) / E^2$$

Para comparar dos medias:

$$n = 2\sigma^2 \times (z_{\alpha} + z_{\beta})^2 / \delta^2$$

Donde δ es la diferencia mínima detectable.

4.8.3 Aplicación práctica: Estudio de prevalencia

Un veterinario necesita estimar la prevalencia de dermatitis atópica en perros de una ciudad.

```

# Parámetros para el cálculo
prevalencia_esperada <- 0.15 # 15% basado en literatura
error_absoluto <- 0.03 # ±3%
confianza <- 0.95
alpha <- 1 - confianza
z <- qnorm(1 - alpha/2)
# Cálculo del tamaño de muestra
n <- (z^2 * prevalencia_esperada * (1 - prevalencia_esperada)) /
error_absoluto^2
n_redondeado <- ceiling(n)
cat("Tamaño de muestra necesario:", n_redondeado, "perros\n")
# Ajuste por población finita
N_poblacion <- 5000 # Población total de perros en la ciudad
n_ajustado <- n / (1 + (n - 1)/N_poblacion)
n_ajustado_redondeado <- ceiling(n_ajustado)
cat("Tamaño ajustado por población finita:", n_ajustado_redondeado, "\n")
# Función para explorar diferentes escenarios
calcular_n <- function(p, error, conf = 0.95) {
  z <- qnorm((1 + conf)/2)
  n <- (z^2 * p * (1 - p)) / error^2
  return(ceiling(n))
}
# Tabla de escenarios
escenarios <- expand.grid(
  prevalencia = c(0.05, 0.10, 0.15, 0.20, 0.25),
  error = c(0.02, 0.03, 0.04, 0.05)
)
escenarios$n <- mapply(calcular_n,
  escenarios$prevalencia,
  escenarios$error)
# Visualización
library(ggplot2)
ggplot(escenarios, aes(x = prevalencia * 100, y = n,
  color = factor(error * 100))) +
  geom_line(size = 1) +
  geom_point(size = 2) +
  labs(title = "Tamaño de Muestra para Estudios de Prevalencia",
    x = "Prevalencia esperada (%)",
    y = "Tamaño de muestra",
    color = "Error (%)") +
  theme_minimal() +
  scale_color_manual(values = c("red", "orange", "blue", "green"),
    labels = c("±2%", "±3%", "±4%", "±5%"))

```

Figura 50. Estudio de prevalencia.

4.8.4 Tamaño de muestra para ensayos clínicos

Un ensayo clínico comparará un nuevo analgésico con el tratamiento estándar en cirugía ortopédica canina.

```

# Parámetros basados en estudios piloto
media_control <- 7.2 # Escala de dolor (0-10)
media_tratamiento <- 5.8 # Reducción esperada
desv_estandar <- 1.8 # Variabilidad común
alpha <- 0.05
potencia <- 0.80
# Cálculo del tamaño del efecto
diferencia <- media_control - media_tratamiento
d_cohen <- diferencia / desv_estandar
cat("Diferencia esperada:", diferencia, "puntos\n")
cat("Tamaño del efecto (d de Cohen):", round(d_cohen, 2), "\n")
# Usando pwr
library(pwr)
resultado_pwr <- pwr.t.test(d = d_cohen,
                           sig.level = alpha,
                           power = potencia,
                           type = "two.sample")
n_por_grupo <- ceiling(resultado_pwr$n)
cat("Animales por grupo:", n_por_grupo, "\n")
# Ajuste por pérdidas esperadas
tasa_perdida <- 0.15 # 15% pérdida esperada
n_ajustado <- ceiling(n_por_grupo / (1 - tasa_perdida))
cat("Con ajuste por pérdidas (15%):", n_ajustado, "por grupo\n")
cat("Total de animales a reclutar:", 2 * n_ajustado, "\n")
# Análisis de sensibilidad
diferencias <- seq(0.5, 2.5, by = 0.1)
ns <- sapply(diferencias, function(dif) {
  d <- dif / desv_estandar
  pwr.t.test(d = d, sig.level = alpha, power = potencia,
             type = "two.sample")$n
})
datos_sens <- data.frame(
  diferencia = diferencias,
  n_por_grupo = ceiling(ns)
)
ggplot(datos_sens, aes(x = diferencia, y = n_por_grupo)) +
  geom_line(size = 1.5) +
  geom_vline(xintercept = 1.4, linetype = "dashed", color = "red") +
  geom_hline(yintercept = n_por_grupo, linetype = "dashed", color = "blue")
+
  annotate("text", x = 1.4, y = 100, label = "Diferencia\nesperada",
          color = "red", hjust = -0.1) +
  labs(title = "Tamaño de Muestra según Diferencia Mínima Detectable",
       x = "Diferencia en escala de dolor",
       y = "n por grupo") +

```

Figura 51. Tamaño de muestra para ensayos clínicos.

4.8.5 Consideraciones especiales en veterinaria

Efecto de Diseño (DEFF) en muestreo por conglomerados

Cuando muestreamos animales agrupados (rebaños, camadas), debemos ajustar por correlación intraclase.

```
# Ejemplo: Muestreo de cerdos en granjas
n_simple <- 384 # Tamaño para muestreo aleatorio simple
ICC <- 0.15 # Correlación intraclase para enfermedad respiratoria
m <- 20 # Animales por granja

# Efecto de diseño
DEFF <- 1 + (m - 1) * ICC
n_ajustado <- n_simple * DEFF
n_granjas <- ceiling(n_ajustado / m)

cat("Efecto de diseño (DEFF):", round(DEFF, 2), "\n")
cat("Animales totales necesarios:", ceiling(n_ajustado), "\n")
cat("Número de granjas:", n_granjas, "\n")
cat("Total real de animales:", n_granjas * m, "\n")

# Visualización del impacto del ICC
ICCs <- seq(0, 0.3, by = 0.01)
DEFFs <- 1 + (m - 1) * ICCs
n_totales <- n_simple * DEFFs

datos_ICC <- data.frame(ICC = ICCs, n_total = n_totales)

ggplot(datos_ICC, aes(x = ICC, y = n_total)) +
  geom_line(size = 1.5, color = "darkgreen") +
  geom_hline(yintercept = n_simple, linetype = "dashed", color = "red") +
  annotate("text", x = 0.25, y = n_simple + 50,
    label = "n sin ajuste", color = "red") +
  labs(title = "Impacto de la Correlación Intraclase en el Tamaño de Muestra",
    x = "Correlación Intraclase (ICC)",
    y = "Tamaño de muestra total") +
  theme_minimal()
```

Figura 52. Efecto de Diseño (DEFF) en muestreo por conglomerados.

Ejercicio 1: Método de ecuación de recursos

Para diseños experimentales con animales, equilibra consideraciones estadísticas y éticas.

```
# Ecuación de recursos: E = N - k
# E debe estar entre 10 y 20
calcular_E <- function(n_total, n_grupos) {
  E <- n_total - n_grupos
  return(E)
}
# Ejemplo: Experimento factorial 2x2
factores <- 2
niveles <- 2
n_tratamientos <- factores * niveles
replicas <- c(3, 4, 5, 6, 7, 8, 10, 12)
resultados_E <- data.frame(
  replicas = replicas,
  n_total = replicas * n_tratamientos,
  E = replicas * n_tratamientos - n_tratamientos
)
resultados_E$recomendacion <- ifelse(
  resultados_E$E < 10, "Muy bajo",
  ifelse(resultados_E$E > 20, "Excesivo", "Adecuado")
)
print(resultados_E)
# Visualización
library(ggplot2)
ggplot(resultados_E, aes(x = replicas, y = E, color = recomendacion)) +
  geom_point(size = 4) +
  geom_line(size = 1) +
  geom_hline(yintercept = c(10, 20), linetype = "dashed", alpha = 0.5) +
  annotate("rect", xmin = -Inf, xmax = Inf, ymin = 10, ymax = 20,
    alpha = 0.2, fill = "green") +
  labs(title = "Ecuación de Recursos para Diseño Factorial 2x2",
    x = "Réplicas por tratamiento",
    y = "E (grados de libertad del error)",
    color = "Recomendación") +
  theme_minimal() +
  scale_color_manual(values = c("Adecuado" = "green",
    "Muy bajo" = "red",
    "Excesivo" = "orange"))
```

Figura 53. Método de ecuación de recursos.

Ejercicio 2: Comparación de tratamientos

Compare la eficacia de dos desparasitantes en la reducción de huevos por gramo (HPG) de heces

```

# Datos pre y post tratamiento
desparasitante_A <- list(
  pre = c(850, 920, 780, 1050, 890, 975, 825, 910, 860, 995),
  post = c(120, 85, 95, 145, 110, 125, 90, 105, 115, 130)
)desparasitante_B <- list(
  pre = c(875, 940, 810, 980, 865, 1020, 845, 895, 925, 970),
  post = c(250, 285, 220, 310, 245, 295, 215, 260, 275, 290)
)# Análisis completo
analizar_eficacia <- function(trat_A, trat_B) {
  # Calcular reducción porcentual
  reduccion_A <- (trat_A$pre - trat_A$post) / trat_A$pre * 100
  reduccion_B <- (trat_B$pre - trat_B$post) / trat_B$pre * 100
  cat("ANÁLISIS DE EFICACIA DE DESPARASITANTES\n")
  cat("=====\n\n")
  # Estadísticas por grupo
  cat("Desparasitante A:\n")
  cat(" Reducción media:", round(mean(reduccion_A), 1), "%\n")
  cat(" IC 95%:", round(t.test(reduccion_A)$conf.int, 1), "%\n\n")
  cat("Desparasitante B:\n")
  cat(" Reducción media:", round(mean(reduccion_B), 1), "%\n")
  cat(" IC 95%:", round(t.test(reduccion_B)$conf.int, 1), "%\n\n")
  # Comparación
  comparacion <- t.test(reduccion_A, reduccion_B, var.equal = FALSE)
  cat("Comparación de eficacia:\n")
  cat(" Diferencia:", round(mean(reduccion_A) - mean(reduccion_B), 1),
"%\n")
  cat(" IC 95% diferencia:", round(comparacion$conf.int, 1), "%\n")
  cat(" Valor p:", round(comparacion$p.value, 4), "%\n")
  cat(" Conclusión:", ifelse(comparacion$p.value < 0.05,
  "Diferencia significativa", "No hay diferencia significativa"), "%\n\n")
  # Visualización
  library(ggplot2) datos_plot <- data.frame(
    Desparasitante = rep(c("A", "B"), each = 10),
    Reduccion = c(reduccion_A, reduccion_B)
  )
  p <- ggplot(datos_plot, aes(x = Desparasitante, y = Reduccion,
    fill = Desparasitante)) +
    geom_boxplot(alpha = 0.7) + geom_point(position = position_jitter(width
= 0.1), size = 2) +
    labs(title = "Comparación de Eficacia de Desparasitantes",
      y = "Reducción en HPG (%)") + theme_minimal() +
    scale_fill_manual(values = c("lightblue", "lightgreen"))
  print(p) return(list(reduccion_A = reduccion_A, reduccion_B =
reduccion_B, test = comparacion))}resultado <-
analizar_eficacia(desparasitante_A, desparasitante_B)

```

Figura 54. Comparación de Tratamientos.

Ejercicio 3: Cálculo de tamaño de muestra

Diseñe un estudio para evaluar la prevalencia de displasia de cadera en una raza canina:

```

# Parámetros del problema
disenar_estudio_displasia <- function(
  prevalencia_esperada = 0.18, # 18% según literatura
  precision = 0.04,           # ±4%
  confianza = 0.95,
  poblacion_total = 2000,     # Perros de la raza en la región
  costo_por_perro = 50        # USD por radiografía
) { cat("DISEÑO DE ESTUDIO: DISPLASIA DE CADERA\n")
  cat("=====\n\n")
  # Cálculo básico  $z <- qnorm((1 + confianza)/2)$   $n\_infinita <- (z^2 *
prevalencia\_esperada * (1 - prevalencia\_esperada)) /$ 
    precision^2 # Corrección por población finita
  n_finita <- n_infinita / (1 + (n_infinita - 1)/poblacion_total)
  # Ajuste por no respuesta (15%)
  tasa_respuesta <- 0.85
  n_ajustada <- n_finita / tasa_respuesta
  # Cálculos de costo
  costo_total <- ceiling(n_ajustada) * costo_por_perro
  # Resultados cat("Tamaño de muestra:\n")
  cat(" Población infinita:", ceiling(n_infinita), "perros\n")
  cat(" Con corrección finita:", ceiling(n_finita), "perros\n")
  cat(" Ajustado por no respuesta:", ceiling(n_ajustada), "perros\n\n")
  cat("Consideraciones prácticas:\n")
  cat(" Población objetivo:", poblacion_total, "perros\n")
  cat(" Proporción a muestrear:",
    round(n_ajustada/poblacion_total*100, 1), "%\n")
  cat(" Costo estimado: $", costo_total, "USD\n\n")
  # Análisis de sensibilidad
  prevalencias <- seq(0.10, 0.30, by = 0.02)
  ns <- sapply(prevalencias, function(p) {
    n <- (z^2 * p * (1 - p)) / precision^2
    n <- n / (1 + (n - 1)/poblacion_total)
    n / tasa_respuesta }) datos_sens <- data.frame( prevalencia = prevalencias
* 100, n = ceiling(ns), costo = ceiling(ns) * costo_por_perro
) # Gráfico de sensibilidad library(ggplot2)
p1 <- ggplot(datos_sens, aes(x = prevalencia, y = n)) + geom_line(size = 1.5, color = "blue") +
  geom_vline(xintercept = prevalencia_esperada * 100,
    linetype = "dashed", color = "red") + labs(title = "Tamaño de Muestra según Prevalencia
Esperada", x = "Prevalencia esperada (%)",
  y = "n necesario") + theme_minimal() p2 <- ggplot(datos_sens, aes(x = prevalencia, y =
costo/1000)) + geom_line(size = 1.5, color = "green") + geom_vline(xintercept =
prevalencia_esperada * 100, linetype = "dashed", color = "red") + labs(title = "Costo del
Estudio",
  x = "Prevalencia esperada (%)", y = "Costo (miles USD)") + theme_minimal()
library(gridExtra) grid.arrange(p1, p2, ncol = 2) return(list(n_final = ceiling(n_ajustada),
costo = costo_total))} # Ejecutar diseño disenar_estudio_displasia()

```

Figura 55. Cálculo de tamaño de muestra.

```

# Gráfico de sensibilidad
library(ggplot2)
p1 <- ggplot(datos_sens, aes(x = prevalencia, y = n)) +
  geom_line(size = 1.5, color = "blue") +
  geom_vline(xintercept = prevalencia_esperada * 100,
             linetype = "dashed", color = "red") +
  labs(title = "Tamaño de Muestra según Prevalencia Esperada",
       x = "Prevalencia esperada (%)",
       y = "n necesario") +
  theme_minimal()

p2 <- ggplot(datos_sens, aes(x = prevalencia, y = costo/1000)) +
  geom_line(size = 1.5, color = "green") +
  geom_vline(xintercept = prevalencia_esperada * 100,
             linetype = "dashed", color = "red") +
  labs(title = "Costo del Estudio",
       x = "Prevalencia esperada (%)",
       y = "Costo (miles USD)") +
  theme_minimal()

library(gridExtra)
grid.arrange(p1, p2, ncol = 2)

return(list(n_final = ceiling(n_ajustada),
           costo = costo_total))
}

# Ejecutar diseño
diseño <- disenar_estudio_displasia()

```

Figura 56. Cálculo de tamaño de muestra.

4.9 Problemas Propuestos

1. **Análisis de Parámetros Sanguíneos:** Un laboratorio veterinario desea establecer valores de referencia para fosfatasa alcalina en cachorros. Analice una muestra de 40 cachorros sanos y construya el intervalo de referencia del 95%.
2. **Ensayo de Vacuna:** Diseñe un estudio para evaluar una nueva vacuna contra parvovirus canino. La vacuna actual tiene 85% de eficacia. ¿Cuántos perros necesita para detectar una mejora del 10% con potencia del 90%?

3. **Estudio Longitudinal:** Evalúe el efecto de un suplemento nutricional en el crecimiento de terneros, con mediciones mensuales durante 6 meses. Considere la correlación entre medidas repetidas.
4. **Meta-análisis:** Combine resultados de 5 estudios sobre prevalencia de leptospirosis en perros urbanos. Calcule la prevalencia pooled y evalúe la heterogeneidad.

4.10 Conclusiones del capítulo

El presente capítulo ha desarrollado los pilares fundamentales de la inferencia estadística aplicada a las ciencias veterinarias, estableciendo un marco metodológico integral que permite al profesional veterinario transitar desde la observación muestral hacia conclusiones poblacionales válidas y confiables. La inferencia estadística no constituye meramente un conjunto de técnicas matemáticas, sino un paradigma de pensamiento crítico que fundamenta la medicina veterinaria basada en evidencia. La estimación puntual y por intervalos proporciona las herramientas básicas para caracterizar parámetros poblacionales a partir de muestras limitadas.

Los intervalos de confianza, en particular, han demostrado su superioridad sobre las estimaciones puntuales al cuantificar explícitamente la incertidumbre inherente al proceso de muestreo. Esta cuantificación resulta esencial en contextos veterinarios donde las decisiones clínicas y epidemiológicas tienen implicaciones directas sobre el bienestar animal, la salud pública y la economía pecuaria. Las pruebas de hipótesis, cuando se aplican apropiadamente, permiten

evaluar afirmaciones científicas con rigor metodológico. Sin embargo, la comprensión de sus limitaciones —particularmente la dependencia del tamaño muestral y la dicotomía artificial entre "significativo" y "no significativo"— resulta crucial para evitar interpretaciones erróneas que han plagado históricamente la literatura veterinaria. La integración de pruebas de hipótesis con intervalos de confianza proporciona una visión más completa, combinando la evaluación de significancia estadística con la magnitud y precisión del efecto.

La comprensión profunda de los errores tipo I y tipo II trasciende el ámbito teórico para convertirse en una consideración práctica fundamental en el diseño de estudios veterinarios. La decisión sobre qué tipo de error minimizar depende crucialmente del contexto: mientras que en la aprobación de nuevos fármacos veterinarios se prioriza el control del error tipo I para proteger la seguridad animal, en programas de detección temprana de enfermedades zoonóticas, minimizar el error tipo II puede ser vital para la salud pública.

El concepto de potencia estadística emerge como un elemento central en la planificación de investigación veterinaria éticamente responsable. Un estudio con potencia insuficiente no solo desperdicia recursos económicos, sino que también viola principios éticos fundamentales al exponer animales a procedimientos experimentales sin probabilidad razonable de detectar efectos clínicamente relevantes. La evidencia presentada de que solo 22% de ensayos clínicos veterinarios reporta cálculos de tamaño muestral y que 43% de estudios con resultados negativos presenta potencia inferior al 50% subraya la necesidad urgente de mejorar los estándares metodológicos en la profesión.

El cálculo apropiado del tamaño de muestra representa la intersección entre consideraciones estadísticas, éticas y prácticas. La implementación del principio de las 3Rs (Reemplazo, Reducción y Refinamiento) en investigación animal requiere una optimización cuidadosa que evite tanto el uso excesivo como insuficiente de animales. Los métodos presentados, desde fórmulas básicas hasta aproximaciones complejas que consideran estructuras jerárquicas y correlaciones intraclase, proporcionan las herramientas necesarias para este balance delicado.

La implementación práctica de los conceptos de inferencia estadística se ha facilitado enormemente mediante el uso de software especializado, particularmente R. Los ejemplos desarrollados a lo largo del capítulo demuestran que la complejidad computacional ya no representa una barrera para la aplicación de métodos estadísticos apropiados. La capacidad de realizar simulaciones, análisis de sensibilidad y visualizaciones interactivas permite una comprensión más profunda de los conceptos abstractos y facilita la comunicación de resultados a audiencias diversas.

CAPÍTULO V

5 PRUEBAS PARAMÉTRICAS

5.1 Objetivos del capítulo

Al finalizar este capítulo, el estudiante será capaz de:

- Comprender los fundamentos teóricos de las pruebas paramétricas y su aplicación en medicina veterinaria
- Seleccionar y aplicar correctamente las pruebas t de Student según el diseño experimental
- Realizar e interpretar análisis de varianza de uno y dos factores
- Ejecutar pruebas post-hoc apropiadas para comparaciones múltiples
- Verificar los supuestos estadísticos y aplicar medidas correctivas cuando sea necesario

5.2 Introducción

Las pruebas paramétricas constituyen herramientas fundamentales en la investigación veterinaria, permitiendo comparar grupos y evaluar hipótesis sobre parámetros poblacionales. Estas pruebas asumen que los datos provienen de distribuciones conocidas, generalmente la distribución normal, y requieren el cumplimiento de ciertos supuestos para su validez. En el contexto veterinario, estas pruebas se aplican

frecuentemente en ensayos clínicos, evaluación de tratamientos, estudios de producción animal y comparación de métodos diagnósticos.

5.3 Prueba t de Student para una muestra

5.3.1 Fundamentos teóricos

La prueba t para una muestra evalúa si la media de una población difiere significativamente de un valor teórico o de referencia. En medicina veterinaria, esta prueba resulta especialmente útil cuando se comparan parámetros fisiológicos con valores de referencia establecidos en la literatura (Arias et al., 2020).

El estadístico t se calcula mediante la fórmula:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Donde:

- $\bar{x}$ = media muestral
- μ_0 = valor hipotético de la media poblacional
- s = desviación estándar muestral
- n = tamaño de muestra

Los grados de libertad (gl) equivalen a $n - 1$.

5.3.2 Hipótesis

- **Hipótesis nula (H_0):** $\mu = \mu_0$ (la media poblacional es igual al valor de referencia)
- **Hipótesis alternativa (H_1):**
 - Bilateral: $\mu \neq \mu_0$
 - Unilateral derecha: $\mu > \mu_0$
 - Unilateral izquierda: $\mu < \mu_0$

5.3.3 Aplicación en veterinaria: Caso práctico

Ejemplo: Un veterinario desea evaluar si el peso promedio de cachorros de raza Labrador Retriever a las 8 semanas en su clínica difiere del estándar internacional de 8.5 kg.

Datos recolectados: Se pesaron 15 cachorros, obteniendo una media de 7.8 kg y una desviación estándar de 0.9 kg.

Procedimiento paso a paso:

1. Establecer las hipótesis:
 - $H_0: \mu = 8.5 \text{ kg}$
 - $H_1: \mu \neq 8.5 \text{ kg}$ (prueba bilateral)
 - Nivel de significancia: $\alpha = 0.05$

2. Calcular el estadístico t:
$$t = \frac{7.8 - 8.5}{0.9/\sqrt{15}} = \frac{-0.7}{0.232} = -3.01$$
3. Determinar grados de libertad: $gl = 15 - 1 = 14$
4. Valor crítico (tabla t, bilateral): $t_{0.025,14} = \pm 2.145$
5. Decisión: Como $|-3.01| > 2.145$, se rechaza H_0
6. Conclusión: Existe evidencia estadística significativa ($p < 0.05$) de que el peso promedio de los cachorros en la clínica difiere del estándar internacional.

5.3.4 Consideraciones prácticas

La interpretación clínica sugiere que los cachorros presentan un peso inferior al esperado, lo cual podría indicar problemas nutricionales o de manejo que requieren investigación adicional. Este hallazgo tiene implicaciones para el protocolo de alimentación y seguimiento del crecimiento en la clínica.

5.4 Prueba t para muestras independientes

5.4.1 Fundamentos teóricos

Esta prueba compara las medias de dos grupos independientes, siendo fundamental en ensayos clínicos veterinarios donde se evalúan tratamientos en grupos separados de animales (Al-kassab, 2022).

El estadístico t para varianzas iguales se calcula como:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Donde s_p es la desviación estándar combinada:

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

5.4.2 Prueba de Levene para homogeneidad de varianzas

Antes de aplicar la prueba t, se debe verificar la igualdad de varianzas mediante la prueba de Levene. Si las varianzas no son homogéneas ($p < 0.05$ en Levene), se utiliza la corrección de Welch (Molina, 2022).

5.5 Prueba t para muestras pareadas

5.5.1 Fundamentos teóricos

La prueba t pareada analiza diferencias cuando las observaciones están relacionadas, como mediciones antes-después en los mismos animales o comparaciones en animales gemelos (Zweifach, 2024).

El estadístico se calcula sobre las diferencias (d):

$$t = \frac{\bar{d}}{s_d / \sqrt{n}}$$

Donde:

- $\bar{d}$ = media de las diferencias

- s_d = desviación estándar de las diferencias
- n = número de pares

5.5.2 Aplicación en veterinaria: Caso práctico

Ejemplo: Evaluación del efecto de un suplemento nutricional sobre los niveles de hemoglobina en perros con anemia leve.

Diseño: Medición de hemoglobina antes y después de 30 días de suplementación en 12 perros.

Tabla 24. Datos (g/dL).

Perro	Antes	Después	Diferencia (d)
1	10.2	12.1	1.9
2	9.8	11.5	1.7
3	10.5	12.8	2.3
4	9.5	11.2	1.7
5	10.1	11.9	1.8
6	9.7	11.8	2.1
7	10.3	12.5	2.2
8	9.9	11.6	1.7
9	10.0	12.2	2.2
10	9.6	11.3	1.7

11	10.4	12.7	2.3
12	9.8	11.9	2.1

Análisis:

1. Calcular estadísticos de las diferencias:
 - $\bar{d} = 1.975 \text{ g/dL}$
 - $s_d = 0.258 \text{ g/dL}$
2. Calcular el estadístico t: $t = \frac{1.975}{0.258/\sqrt{12}} = \frac{1.975}{0.0745} = 26.51$
3. Grados de libertad: $gl = 11$
4. Valor crítico ($\alpha = 0.05$, bilateral): $t_{0.025,11} = \pm 2.201$
5. Valor $p < 0.001$
6. Intervalo de confianza al 95%: $1.975 \pm 2.201 \times 0.0745 = [1.81, 2.14]$

5.5.3 Interpretación y relevancia clínica

El suplemento nutricional produce un incremento significativo en los niveles de hemoglobina ($p < 0.001$), con un aumento promedio de 1.98 g/dL (IC 95%: 1.81-2.14). Este incremento es clínicamente relevante, representando una mejora aproximada del 20% en los valores basales.

5.4 Análisis de varianza (ANOVA) de un factor

5.5.4 Fundamentos teóricos

El ANOVA de un factor extiende la prueba t para comparar simultáneamente las medias de tres o más grupos independientes, controlando el error tipo I que aumentaría al realizar múltiples pruebas t (Zhao et al., 2021).

El estadístico F se calcula como:

$$F = \frac{\text{Varianza entre grupos}}{\text{Varianza dentro de grupos}} = \frac{CME}{CMD}$$

Tabla 25. La tabla ANOVA.

Fuente de Variación	SC	gl	CM	F	p
Entre grupos	SCE	k-1	CME = SCE/(k-1)	CME/CMD	
Dentro de grupos	SCD	N-k	CMD = SCD/(N-k)		
Total	SCT	N-1			

5.5.5 Aplicación en veterinaria: Caso práctico

Ejemplo: Comparación de tres protocolos anestésicos en la recuperación postquirúrgica de gatos.

Grupos:

- Protocolo A: Ketamina-Xilacina (n=15)
- Protocolo B: Propofol-Isoflurano (n=15)

- Protocolo C: Dexmedetomidina-Ketamina (n=15)
- Variable: Tiempo de recuperación (minutos)

Tabla 26. Datos resumidos.

Protocolo	n	Media	Desv. Est.	Varianza
A	15	45.3	8.2	67.24
B	15	32.1	6.5	42.25
C	15	38.7	7.1	50.41

5.5.6 Cálculos del ANOVA

1. Media global: $\bar{x} = 38.7$ minutos
2. Suma de cuadrados entre grupos: $SSCE = 15[(45.3-38.7)^2 + (32.1-38.7)^2 + (38.7-38.7)^2] = 1306.8$
3. Suma de cuadrados dentro de grupos: $SSCD = 14(67.24 + 42.25 + 50.41) = 2238.6$

Tabla 27. Tabla ANOVA.

Fuente	SC	gl	CM	F	p
Entre protocolos	1306.8	2	653.4	12.26	<0.001
Error	2238.6	42	53.3		
Total	3545.4	44			

5.4.3 Interpretación

El análisis revela diferencias significativas entre los protocolos anestésicos ($F_{(2,42)} = 12.26$, $p < 0.001$). Sin embargo, el ANOVA no indica qué protocolos difieren entre sí, requiriendo pruebas post-hoc para identificar las diferencias específicas.

5.6 ANOVA de dos factores

5.6.1 Fundamentos teóricos

El ANOVA de dos factores evalúa el efecto de dos variables independientes (factores) y su posible interacción sobre una variable dependiente continua (Hothorn, 2022). En veterinaria, permite estudiar efectos combinados de tratamientos y condiciones ambientales, sexo y dieta, entre otros.

El modelo estadístico es:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

Donde:

- μ = media general
- α_i = efecto del nivel i del factor A
- β_j = efecto del nivel j del factor B
- $(\alpha\beta)_{ij}$ = efecto de la interacción

- ϵ_{ijk} = error aleatorio

5.6.2 Aplicación en veterinaria: Caso práctico

Ejemplo: Estudio del efecto del sexo y la dieta sobre la ganancia de peso en conejos.

Diseño factorial 2×3:

- Factor A (Sexo): Macho, Hembra
- Factor B (Dieta): Estándar, Alta proteína, Alta energía
- $n = 10$ animales por combinación
- Variable: Ganancia de peso (g) en 30 días

Tabla 28. Datos resumidos.

Sexo	Dieta	Media	Desv. Est.
Macho	Estándar	450	35
Macho	Alta proteína	520	40
Macho	Alta energía	495	38
Hembra	Estándar	420	32
Hembra	Alta proteína	480	36
Hembra	Alta energía	510	42

Tabla 29. Tabla ANOVA de dos factores.

Fuente	SC	gl	CM	F	p	η^2 parcial
Sexo	4083.3	1	4083.3	3.12	0.082	0.05
Dieta	28816.7	2	14408.3	11.01	<0.001	0.29
Sexo $\times$ Dieta	3616.7	2	1808.3	1.38	0.259	0.05
Error	70800.0	54	1311.1			
Total	107316.7	59				

5.6.3 Interpretación de resultados

1. **Efecto principal del sexo:** No significativo ($p = 0.082$), aunque muestra tendencia
2. **Efecto principal de la dieta:** Altamente significativo ($p < 0.001$)
3. **Interacción sexo $\times$ dieta:** No significativa ($p = 0.259$)

La dieta tiene un efecto sustancial sobre la ganancia de peso ($\eta^2 = 0.29$), explicando el 29% de la variabilidad. La ausencia de interacción significativa indica que el efecto de la dieta es similar en ambos sexos.

5.6.4 Gráfico de interacción

El análisis visual mediante gráficos de interacción complementa la interpretación estadística. Las líneas aproximadamente paralelas confirman la ausencia de interacción significativa.

5.7 Pruebas Post-Hoc

5.7.1 Fundamentos y selección

Las pruebas post-hoc realizan comparaciones múltiples después de un ANOVA significativo, controlando el error tipo I global. La elección depende de varios factores (Dugard et al., 2010):

Criterios de selección:

- **Tamaños de muestra iguales:** Tukey HSD, Student-Newman-Keuls
- **Tamaños de muestra desiguales:** Tukey-Kramer, Games-Howell
- **Comparaciones planificadas:** Contrastes ortogonales, Dunnett
- **Control conservador:** Bonferroni, Scheffé

5.7.2 Prueba de Tukey HSD

La diferencia honestamente significativa (HSD) de Tukey es apropiada cuando se comparan todos los pares posibles con tamaños de muestra iguales.

$$HSD = q_{\alpha,k,g} \sqrt{\frac{CME}{n}}$$

5.7.3 Aplicación práctica

Continuando con el ejemplo de protocolos anestésicos (sección 5.4):

Cálculo de Tukey HSD:

1. $CME = 53.3$, $n = 15$, $k = 3$, $gl = 42$
2. $q_{0.05,3,42} = 3.44$ (tabla de rangos studentizados)
3. $HSD = 3.44 \times \sqrt{(53.3/15)} = 6.48$ minutos

Tabla 30. Comparaciones pareadas.

Comparación	Diferencia	> HSD	Conclusión
A vs B	$45.3 - 32.1 = 13.2$	Sí	Significativo
A vs C	$45.3 - 38.7 = 6.6$	Sí	Significativo
B vs C	$32.1 - 38.7 = -6.6$	Sí	Significativo

5.7.4 Prueba de Dunnett

Utilizada cuando se comparan varios tratamientos contra un control. En veterinaria, es común en estudios donde nuevos fármacos se comparan con un tratamiento estándar (Quicke et al., 2021).

Ejemplo: Comparación de tres nuevos antibióticos con amoxicilina (control) en el tratamiento de mastitis bovina.

El valor crítico de Dunnett es menor que el de Tukey, proporcionando mayor potencia para detectar diferencias con el control.

5.7.5 Corrección de Bonferroni

Método conservador que ajusta el nivel de significancia:

$$\alpha_{ajustado} = \frac{\alpha}{m}$$

Donde m es el número de comparaciones.

Para 6 comparaciones con $\alpha = 0.05$:

$$\alpha_{ajustado} = \frac{0.05}{6} = 0.0083$$

5.8 Supuestos y verificación

5.8.1 Supuestos de las pruebas paramétricas

Las pruebas paramétricas requieren el cumplimiento de tres supuestos principales (Backhaus et al., 2021):

1. **Normalidad:** Los datos provienen de poblaciones con distribución normal
2. **Homogeneidad de varianzas:** Las varianzas poblacionales son iguales (homocedasticidad)
3. **Independencia:** Las observaciones son independientes entre sí

5.8.2 Verificación de normalidad

Métodos gráficos:

- Histogramas con curva normal superpuesta
- Gráficos Q-Q (cuantil-cuantil)

- Gráficos de probabilidad normal

Pruebas estadísticas:

- Shapiro-Wilk: Recomendada para $n < 50$
- Kolmogorov-Smirnov: Para muestras grandes
- Anderson-Darling: Sensible a desviaciones en las colas

```
# Datos de glucosa en sangre de 30 perros
glucosa <- c(92, 95, 88, 103, 97, 91, 99, 102, 94, 96,
            89, 101, 93, 98, 100, 87, 104, 95, 92, 99,
            91, 96, 102, 88, 97, 94, 103, 90, 98, 95)

# Prueba de Shapiro-Wilk
shapiro.test(glucosa)
# W = 0.9734, p-value = 0.6327

# Gráfico Q-Q
qqnorm(glucosa)
qqline(glucosa)
```

Figura 57. Ejemplo de aplicación en R.

Interpretación: Con $p = 0.633 > 0.05$, no se rechaza la hipótesis de normalidad.

5.8.3 Verificación de homogeneidad de varianzas

Prueba de Levene: Robusta ante desviaciones de la normalidad, utiliza las desviaciones absolutas respecto a la mediana.

Prueba de Bartlett: Más sensible, pero requiere normalidad estricta.

Prueba de Brown-Forsythe: Modificación de Levene usando la mediana, más robusta.

Ejemplo práctico:

```
# Comparación de varianzas en tres grupos de tratamiento  
library(car)  
leveneTest(peso ~ tratamiento, data = datos_veterinarios)  
# F = 1.823, p = 0.172
```

5.8.4 Verificación de independencia

La independencia se garantiza principalmente mediante:

- Aleatorización adecuada
- Diseño experimental apropiado
- Evitar mediciones repetidas no planificadas
- Control de efectos de camada o parentesco

Prueba de Durbin-Watson: Detecta autocorrelación en residuos

```
# Para datos ordenados temporalmente  
durbinWatsonTest(modelo_lm)  
# DW = 1.95, p = 0.418
```

5.8.5 Transformaciones de Datos

Cuando los supuestos no se cumplen, las transformaciones pueden normalizar la distribución y estabilizar varianzas:

Transformaciones comunes en veterinaria:

1. **Logarítmica [$\log(x)$]:** Para datos con sesgo positivo (ej. conteos parasitarios)
2. **Raíz cuadrada [$\sqrt{x}$]:** Para datos de conteo (ej. recuento bacteriano)
3. **Arcoseno [$\arcsin(\sqrt{p})$]:** Para proporciones (ej. tasas de mortalidad)
4. **Box-Cox:** Transformación óptima determinada por máxima verosimilitud

Ejemplo de transformación logarítmica:

```
# Conteo de parásitos con distribución sesgada
parasitos <- c(10, 15, 8, 120, 95, 12, 18, 150, 9, 11,
              85, 13, 16, 110, 14, 90, 17, 19, 88, 20)

# Verificar normalidad original
shapiro.test(parasitos) # p = 0.0001

# Aplicar transformación logarítmica
log_parasitos <- log10(parasitos + 1) # +1 para evitar log(0)

# Verificar normalidad después de transformación
shapiro.test(log_parasitos) # p = 0.084
```

Figura 58. Ejemplo de transformación logarítmica.

5.8.6 Análisis de residuos

El análisis de residuos es fundamental para validar el modelo:

Tipos de gráficos de residuos:

1. **Residuos vs valores ajustados:** Detecta heterocedasticidad
2. **Q-Q de residuos:** Evalúa normalidad de residuos
3. **Residuos vs orden:** Detecta dependencia temporal
4. **Residuos estandarizados:** Identifica valores atípicos

Criterios de evaluación:

- Residuos distribuidos aleatoriamente alrededor de cero
- Varianza constante (banda horizontal uniforme)
- Ausencia de patrones sistemáticos
- Residuos estandarizados dentro de ± 3

5.8.7 Manejo de violaciones de supuestos

Estrategias según el tipo de violación:

1. Violación de normalidad:

- Transformación de datos
- Pruebas no paramétricas
- Métodos robustos o bootstrap

2. Violación de homocedasticidad:

- Transformación estabilizadora de varianza
- Prueba t de Welch (para dos grupos)
- ANOVA de Welch o Brown-Forsythe

3. Violación de independencia:

- Modelos mixtos para datos correlacionados
- Análisis de medidas repetidas
- Corrección de errores estándar

5.8.8 Robustez de las pruebas paramétricas

Las pruebas t y ANOVA muestran cierta robustez ante violaciones moderadas:

- Robustez ante no normalidad: Con $n \geq 30$ por grupo, el teorema del límite central proporciona protección
- Robustez ante heterocedasticidad: Si los tamaños de muestra son iguales y la razón de varianzas $< 3:1$
- No robustez ante dependencia: La violación de independencia invalida los resultados

5.9 Ejercicios propuestos

Ejercicio 5.1 (Nivel Básico)

Un veterinario registra los niveles de cortisol ($\mu\text{g/dL}$) en 10 perros antes de una cirugía: 2.8, 3.1, 2.9, 3.3, 2.7, 3.0, 2.8, 3.2, 2.9, 3.1. El valor de referencia normal es 2.5 $\mu\text{g/dL}$. ¿Existe evidencia de que los niveles preoperatorios son superiores al valor normal? Use $\alpha = 0.05$.

Ejercicio 5.2 (Nivel Intermedio)

Se evalúan dos métodos de inseminación artificial en vacas lecheras. El método A ($n=25$) tiene una tasa de concepción media del 65% ($s=12\%$), mientras que el método B ($n=30$) presenta 58% ($s=15\%$). Realice el análisis completo incluyendo verificación de supuestos e intervalo de confianza para la diferencia.

Ejercicio 5.3 (Nivel Avanzado)

Diseñe y analice un experimento factorial 2×3 para evaluar el efecto del tipo de alojamiento (individual vs grupal) y tres niveles de proteína en la dieta (14%, 16%, 18%) sobre la conversión alimenticia en cerdos en crecimiento. Incluya: a) Tabla ANOVA completa b) Análisis de interacción c) Pruebas post-hoc apropiadas d) Verificación de supuestos e) Interpretación práctica para producción porcina

5.10 Problemas resueltos

5.10.1 Problema 1: Evaluación de un tratamiento para artritis canina

Contexto: Se evalúa la eficacia de un nuevo AINE en perros con osteoartritis midiendo el score de claudicación (0-10) antes y después del tratamiento en 8 pacientes.

Tabla 31. Datos.

Paciente	Antes	Después
1	7	4
2	8	5
3	6	3
4	9	6
5	7	5
6	8	4
7	6	4
8	7	5

Solución completa:

Solución completa:

1. Hipótesis:

- $H_0: \mu_d = 0$ (no hay diferencia)
- $H_1: \mu_d > 0$ (mejora en el score)

2. Cálculo de diferencias:

- $d: 3, 3, 3, 3, 2, 4, 2, 2$
- $\bar{d} = 2.75$

- $sd = 0.707$
- 3. **Estadístico t:** $t = 2.75 / (0.707 / \sqrt{8}) = 11.0$
- 4. **Decisión:**
 - $gl = 7$, $t_{0.05,7} = 1.895$
 - $11.0 > 1.895$, se rechaza H_0
- 5. **Conclusión:** El tratamiento reduce significativamente el score de claudicación ($p < 0.001$), con una mejora promedio de 2.75 puntos (IC 95%: 2.16-3.34).

5.10.2 Problema 2: Comparación de tres métodos de conservación de semen equino

Contexto: Se evalúa la motilidad espermática (%) después de 48 horas de conservación usando tres diluyentes diferentes en semen de 18 sementales (6 por grupo).

Tabla 32. Datos.

Diluyente A	Diluyente B	Diluyente C
72	65	58
75	68	61
71	64	57
73	67	60

74	66	59
76	69	62

Solución completa:

1. Estadísticos descriptivos:

- Grupo A: $\bar{x}_1 = 73.5$, $s_1 = 1.87$
- Grupo B: $\bar{x}_2 = 66.5$, $s_2 = 1.87$
- Grupo C: $\bar{x}_3 = 59.5$, $s_3 = 1.87$
- Media global: $\bar{x}_{..} = 66.5$

2. Verificación de supuestos:

```
# Prueba de normalidad por grupo
shapiro.test(grupoA) # p = 0.91
shapiro.test(grupoB) # p = 0.89
shapiro.test(grupoC) # p = 0.92

# Prueba de homogeneidad de varianzas
bartlett.test(motilidad ~ diluyente) # p = 0.99
```

Figura 59. Verificación de supuestos.

3. Cálculo del ANOVA:

- $SCE = 6[(73.5-66.5)^2 + (66.5-66.5)^2 + (59.5-66.5)^2] = 588$
- $SCD = 5(3.5 + 3.5 + 3.5) = 52.5$
- $SCT = 640.5$

Tabla 33. Tabla ANOVA.

Fuente	SC	gl	CM	F	p
Entre diluyentes	588.0	2	294.0	84.0	<0.001
Error	52.5	15	3.5		
Total	640.5	17			

4. Pruebas post-hoc (Tukey):

○ $HSD = q_{0.05,3,15} \times \sqrt{(3.5/6)} = 3.67 \times 0.764 = 2.80$

Tabla 34. Pruebas post-hoc (Tukey).

Comparación	Diferencia	Conclusión
A vs B	7.0	Significativo
A vs C	14.0	Significativo
B vs C	7.0	Significativo

5. Interpretación: Los tres diluyentes difieren significativamente en su capacidad de preservar la motilidad espermática. El diluyente A es superior, manteniendo 73.5% de motilidad, seguido por B (66.5%) y C (59.5%). Todas las diferencias son estadística y clínicamente significativas.

5.10.3 Problema 3: Interacción entre raza y sistema de alimentación en ovinos

Contexto: Estudio factorial 2×2 evaluando ganancia diaria de peso (g/día) en corderos.

Factores:

- Raza: Dorper vs Katahdin
- Sistema: Pastoreo vs Estabulación
- $n = 8$ corderos por combinación

Tabla 35. Datos resumidos.

Raza	Sistema	Media	DE
Dorper	Pastoreo	185	15
Dorper	Estabulación	245	18
Katahdin	Pastoreo	195	16
Katahdin	Estabulación	220	17

Solución completa:

1. Modelo y sumas de cuadrados:

- $SC_{\text{Raza}} = 16[(190-211.25)^2 + (207.5-211.25)^2] = 7,256.25$
- $SC_{\text{Sistema}} = 16[(190-211.25)^2 + (232.5-211.25)^2] = 14,756.25$
- $SC_{\text{Interacción}} = 8[(185-190-210+211.25)^2 + \dots] = 2,256.25$

- SC_Error = Calculado de las varianzas dentro de grupos = 8,512.00

Tabla 36. Tabla ANOVA factorial.

Fuente	SC	gl	CM	F	p	η^2
Raza	7,256.25	1	7,256.25	24.38	<0.001	0.45
Sistema	14,756.25	1	14,756.25	49.58	<0.001	0.63
Raza $\times$ Sistema	2,256.25	1	2,256.25	7.58	0.010	0.20
Error	8,512.00	28	297.54			
Total	32,780.75	31				

2. Análisis de efectos simples (debido a interacción significativa):

Efecto del sistema en cada raza:

- Dorper: $245 - 185 = 60$ g/día ($t = 9.8$, $p < 0.001$)
- Katahdin: $220 - 195 = 25$ g/día ($t = 4.1$, $p < 0.001$)

Efecto de la raza en cada sistema:

- Pastoreo: $195 - 185 = 10$ g/día ($t = 1.6$, $p = 0.12$)
- Estabulación: $245 - 220 = 25$ g/día ($t = 4.1$, $p < 0.001$)

- ### 3. Interpretación de la interacción:
- La interacción significativa indica que el beneficio de la estabulación es mayor en corderos Dorper (60 g/día adicionales) que en Katahdin (25 g/día adicionales). Esto sugiere que los Dorper responden mejor a sistemas intensivos.

5.11 Actividades de Autoevaluación

Sección A: Preguntas Conceptuales

1. **¿Cuándo es apropiado usar la corrección de Welch en lugar de la prueba t tradicional?** *Respuesta:* La corrección de Welch debe utilizarse cuando la prueba de Levene indica heterogeneidad de varianzas ($p < 0.05$) y los tamaños de muestra son desiguales. Esta corrección ajusta los grados de libertad para compensar la violación del supuesto de homocedasticidad.
2. **Explique por qué realizar múltiples pruebas t aumenta el error tipo I y cómo el ANOVA resuelve este problema.** *Respuesta:* Al realizar múltiples comparaciones, la probabilidad de cometer al menos un error tipo I aumenta según: $\alpha_{\text{familiar}} = 1 - (1 - \alpha)^c$, donde c es el número de comparaciones. Para 6 comparaciones con $\alpha = 0.05$, el error familiarmente sería 0.26. El ANOVA controla este error realizando una prueba global única.
3. **¿Qué indica una interacción significativa en un ANOVA factorial y cómo afecta la interpretación de los efectos principales?** *Respuesta:* Una interacción significativa indica que el efecto de un factor depende del nivel del otro factor. Cuando existe interacción, los efectos principales pueden ser engañosos y deben interpretarse con cautela, siendo necesario analizar los efectos simples para cada nivel del otro factor.

Sección B: Análisis de Casos

Caso 1: Evaluación de Vacunas

Un investigador evalúa la respuesta inmune (título de anticuerpos) a tres vacunas contra parvovirus canino. Los datos son:

- Vacuna A (n=12): $\bar{x} = 1:320$, $s = 80$
- Vacuna B (n=10): $\bar{x} = 1:256$, $s = 64$
- Vacuna C (n=11): $\bar{x} = 1:384$, $s = 96$

Preguntas: a) ¿Qué prueba estadística es apropiada? b) Si el ANOVA es significativo, ¿qué prueba post-hoc recomendaría y por qué? c) ¿Qué transformación podría ser necesaria para estos datos?

Respuestas: a) ANOVA de un factor, pero con precaución por tamaños desiguales b) Tukey-Kramer o Games-Howell, debido a los tamaños de muestra desiguales c) Transformación logarítmica, ya que los títulos de anticuerpos típicamente siguen distribución log-normal

Sección C: Ejercicios de Aplicación con Software

```

# Datos de un estudio de nutrición en pollos de engorde
# Factor: Nivel de lisina en dieta (0.9%, 1.1%, 1.3%)
# Variable respuesta: Ganancia de peso (g) a los 21 días
library(tidyverse)
library(car)
library(agricolae)
# Crear conjunto de datos
set.seed(123)
datos <- data.frame(
  lisina = factor(rep(c("0.9", "1.1", "1.3"), each = 20)),
  peso = c(rnorm(20, 620, 40), # 0.9%
           rnorm(20, 680, 35), # 1.1%
           rnorm(20, 710, 38)) # 1.3%
)
# 1. Estadísticas descriptivas
datos %>%
  group_by(lisina) %>%
  summarise(
    n = n(),
    media = mean(peso),
    de = sd(peso),
    error_std = de/sqrt(n),
    IC_inf = media - qt(0.975, n-1)*error_std,
    IC_sup = media + qt(0.975, n-1)*error_std
  )
# 2. Verificación de supuestos
# Normalidad por grupo
by(datos$peso, datos$lisina, shapiro.test)
# Homogeneidad de varianzas
leveneTest(peso ~ lisina, data = datos)
# 3. ANOVA
modelo <- aov(peso ~ lisina, data = datos)
summary(modelo)
# 4. Diagnóstico del modelo
par(mfrow=c(2,2))
plot(modelo)
# 5. Pruebas post-hoc
TukeyHSD(modelo)
plot(TukeyHSD(modelo))
# 6. Tamaño del efecto
library(effects)
omega_squared(modelo)

```

Figura 60. Ejercicio con R.

5.12 Conclusiones

Las pruebas paramétricas constituyen herramientas estadísticas fundamentales en la investigación veterinaria, proporcionando métodos rigurosos para la comparación de grupos y la evaluación de hipótesis sobre parámetros poblacionales. A través del desarrollo de este capítulo, se han presentado los fundamentos teóricos y aplicaciones prácticas de las principales pruebas paramétricas, desde la prueba t de Student hasta el análisis de varianza factorial, con énfasis particular en su implementación en contextos veterinarios específicos.

La comprensión integral de las pruebas paramétricas requiere reconocer su interdependencia conceptual. La prueba t de Student, en sus tres variantes (una muestra, muestras independientes y muestras pareadas), establece la base para comparaciones entre dos grupos o condiciones. El análisis de varianza extiende estos principios para permitir comparaciones múltiples simultáneas, controlando eficazmente el error tipo I que se acumularía con pruebas t repetidas. Esta progresión lógica desde comparaciones simples hasta diseños factoriales complejos refleja la evolución natural de las preguntas de investigación en medicina veterinaria.

Los casos prácticos presentados demuestran la versatilidad de estas herramientas en diversos contextos veterinarios: desde la evaluación de parámetros fisiológicos en cachorros hasta la comparación de protocolos anestésicos, pasando por estudios de producción animal y ensayos clínicos. Cada aplicación ilustra no solo el procedimiento

estadístico, sino también la interpretación clínica y las implicaciones prácticas de los resultados.

Un aspecto crítico enfatizado a lo largo del capítulo es la verificación rigurosa de los supuestos subyacentes: normalidad, homogeneidad de varianzas e independencia. La validez de las inferencias estadísticas depende fundamentalmente del cumplimiento de estos supuestos. El capítulo proporciona herramientas diagnósticas múltiples, desde métodos gráficos intuitivos hasta pruebas formales como Shapiro-Wilk y Levene, permitiendo al investigador veterinario evaluar sistemáticamente la idoneidad de las pruebas paramétricas para sus datos.

Cuando los supuestos no se cumplen, las estrategias de remediación presentadas transformaciones de datos, pruebas alternativas robustas y métodos no paramétricos ofrecen soluciones prácticas que mantienen la integridad del análisis estadístico. La transformación logarítmica para datos de conteos parasitarios o la transformación arcoseno para proporciones representan aplicaciones específicas comunes en la práctica veterinaria.

El capítulo subraya consistentemente la distinción entre significancia estadística y relevancia clínica. Un resultado estadísticamente significativo no necesariamente implica importancia práctica en el contexto veterinario. Los tamaños del efecto, intervalos de confianza y medidas como η^2 cuadrado parcial proporcionan información complementaria crucial para evaluar la magnitud y relevancia práctica

de los hallazgos. Esta perspectiva dual estadística y clínica es esencial para la toma de decisiones informadas en medicina veterinaria.

El análisis de varianza factorial introduce conceptos avanzados como efectos principales e interacciones, fundamentales para comprender fenómenos complejos en sistemas biológicos. La capacidad de detectar y analizar interacciones permite descubrir relaciones sinérgicas o antagonistas entre factores, como se ilustró en el ejemplo de la interacción entre raza y sistema de alimentación en ovinos. Estos hallazgos tienen implicaciones directas para la optimización de sistemas de producción y protocolos clínicos.

Las pruebas post-hoc presentadas Tukey HSD, Dunnett, Bonferroni ofrecen estrategias diferenciadas para comparaciones múltiples según el diseño experimental y los objetivos de investigación. La selección apropiada de estas pruebas, considerando factores como tamaños de muestra, número de comparaciones y control del error tipo I, refleja la sofisticación requerida en la investigación veterinaria moderna.

CAPÍTULO VI

6 PRUEBAS NO PARAMÉTRICAS EN CIENCIAS VETERINARIAS

6.1 Objetivos del capítulo

Al finalizar este capítulo, el estudiante será capaz de:

1. Comprender los fundamentos teóricos y supuestos de las pruebas no paramétricas
2. Identificar situaciones en medicina veterinaria donde se requiere el uso de estadística no paramétrica
3. Aplicar correctamente las pruebas no paramétricas más utilizadas en investigación veterinaria
4. Interpretar los resultados obtenidos en el contexto clínico y zootécnico
5. Seleccionar la prueba estadística apropiada según el diseño experimental y naturaleza de los datos

6.2 Introducción

Las pruebas no paramétricas constituyen un conjunto de herramientas estadísticas fundamentales en la investigación veterinaria, particularmente cuando los datos no cumplen con los supuestos requeridos por las pruebas paramétricas tradicionales. En la práctica clínica veterinaria, frecuentemente se encuentran situaciones donde las variables presentan distribuciones asimétricas, tamaños muestrales

pequeños o escalas de medición ordinales, condiciones que hacen inadecuado el uso de pruebas como la t de Student o el ANOVA.

El presente capítulo aborda las principales pruebas no paramétricas utilizadas en ciencias veterinarias, proporcionando una comprensión profunda de sus fundamentos teóricos, aplicaciones prácticas y limitaciones. Se enfatiza la importancia de estas técnicas en el análisis de datos clínicos, experimentales y epidemiológicos en medicina veterinaria y zootecnia.

6.3 Prueba de Mann-Whitney U

6.3.1 Fundamentos teóricos

La prueba de Mann-Whitney U, también conocida como prueba de Wilcoxon de rangos sumados, representa la alternativa no paramétrica a la prueba t de Student para muestras independientes. Esta prueba evalúa si dos muestras independientes provienen de poblaciones con la misma distribución, sin asumir normalidad en los datos (Molina, 2023).

La hipótesis nula (H_0) establece que las dos poblaciones son idénticas en su distribución, mientras que la hipótesis alternativa (H_1) puede ser unilateral o bilateral, indicando diferencias en la tendencia central de las distribuciones.

6.3.2 Procedimiento de cálculo

El procedimiento para aplicar la prueba de Mann-Whitney U sigue los siguientes pasos:

1. **Combinación y ordenamiento:** Se combinan las observaciones de ambos grupos y se ordenan de menor a mayor
2. **Asignación de rangos:** Se asignan rangos a todas las observaciones, promediando los rangos en caso de empates
3. **Suma de rangos:** Se calcula la suma de rangos para cada grupo (R_1 y R_2)
4. **Cálculo del estadístico U:** $U_1 = n_1 \times n_2 + [n_1(n_1 + 1)]/2 - R_1$ $U_2 = n_1 \times n_2 + [n_2(n_2 + 1)]/2 - R_2$ Donde n_1 y n_2 son los tamaños muestrales de cada grupo.
5. **Estadístico de prueba:** Se utiliza el menor valor entre U_1 y U_2

6.3.3 Aplicación en medicina veterinaria

Caso Práctico: Evaluación de tratamiento antiparasitario en bovinos

Un médico veterinario investiga la eficacia de un nuevo antiparasitario comparado con el tratamiento convencional. Se evalúa la reducción en el conteo de huevos por gramo de heces (HPG) en 20 bovinos, asignados aleatoriamente a dos grupos (Kumamoto et al., 2021).

Tabla 37. Reducción en HPG según tratamiento.

Tratamiento Nuevo (n=10)	Tratamiento Control (n=10)
85, 92, 78, 95, 88	72, 68, 75, 70, 65
90, 82, 87, 93, 89	73, 69, 71, 67, 74

Análisis paso a paso:

1. Ordenamiento combinado y asignación de rangos
2. Cálculo de sumas de rangos: $R_1 = 155$, $R_2 = 55$
3. Cálculo de U: $U_1 = 100$, $U_2 = 0$
4. Valor crítico ($\alpha = 0.05$, bilateral): 23
5. Decisión: Como $U = 0 < 23$, se rechaza H_0

Interpretación: Existe evidencia estadística significativa ($p < 0.05$) de que el nuevo tratamiento antiparasitario produce una mayor reducción en el conteo de HPG comparado con el tratamiento convencional.

6.3.4 Implementación en Software Estadístico

```
# Datos del ejemplo
nuevo <- c(85, 92, 78, 95, 88, 90, 82, 87, 93, 89)
control <- c(72, 68, 75, 70, 65, 73, 69, 71, 67, 74)

# Prueba de Mann-Whitney
wilcox.test(nuevo, control, alternative = "two.sided")

# Visualización
boxplot(nuevo, control, names = c("Nuevo", "Control"),
        ylab = "Reducción HPG", col = c("lightblue", "lightgreen"))
```

Figura 61. Código en R.

6.4 Prueba de Wilcoxon para muestras pareadas

6.4.1 Fundamentos teóricos

La prueba de Wilcoxon para muestras pareadas constituye la alternativa no paramétrica a la prueba t pareada. Se utiliza cuando las observaciones están naturalmente pareadas o cuando el mismo sujeto es medido en dos condiciones diferentes (Astuhuaman et al., 2018). En investigación

veterinaria, esta situación es común en estudios de antes-después o cuando se comparan dos tratamientos en el mismo animal.

6.4.2 Supuestos y condiciones

La prueba requiere que:

- Las diferencias entre pares sean continuas y simétricamente distribuidas alrededor de la mediana
- Las observaciones dentro de cada par estén relacionadas
- Los pares sean independientes entre sí

6.4.3 Procedimiento de cálculo

1. **Cálculo de diferencias:** $D_i = X_i - Y_i$ para cada par
2. **Eliminación de diferencias nulas:** Se excluyen los pares donde $D_i = 0$
3. **Rangos absolutos:** Se ordenan $|D_i|$ y se asignan rangos
4. **Signos de rangos:** Se mantiene el signo original de cada diferencia
5. **Suma de rangos:** T^+ (suma de rangos positivos) y T^- (suma de rangos negativos)
6. **Estadístico de prueba:** $T = \min(T^+, T^-)$

6.4.4 Aplicación práctica

Caso clínico: Evaluación de glucemia en perros diabéticos

Se evalúa el efecto de un nuevo protocolo de insulino-terapia midiendo la glucemia (mg/dL) en 12 perros diabéticos antes y después del tratamiento.

Tabla 38. Niveles de glucemia pre y post-tratamiento.

Paciente	Pre-tratamiento	Post-tratamiento	Diferencia
1	285	145	140
2	310	168	142
3	295	155	140
4	320	175	145
5	275	142	133
6	305	160	145
7	290	148	142
8	315	170	145
9	280	138	142
10	300	158	142
11	325	180	145
12	270	135	135

Análisis estadístico:

- $T^+ = 78$ (todas las diferencias son positivas)
- $T^- = 0$
- n efectivo = 12
- Valor crítico ($\alpha = 0.05$, bilateral) = 13
- Decisión: Como $T = 0 < 13$, se rechaza H_0

Interpretación veterinaria: El nuevo protocolo de insulino terapia produce una reducción significativa ($p < 0.05$) en los niveles de glucemia en perros diabéticos.

6.5 Prueba de Kruskal-Wallis

6.5.1 Fundamentos teóricos

La prueba de Kruskal-Wallis representa la extensión de la prueba de Mann-Whitney para comparar tres o más grupos independientes, constituyendo la alternativa no paramétrica al ANOVA de una vía (Burbano & Pijal, 2021). Esta prueba evalúa si k muestras independientes provienen de la misma población o de poblaciones con medianas iguales.

6.5.2 Estadístico de prueba

El estadístico H de Kruskal-Wallis se calcula mediante:

$$H = [12/N(N+1)] \times \sum (R_i^2/n_i) - 3(N+1)$$

Donde:

- N = tamaño total de la muestra
- k = número de grupos
- n_i = tamaño del grupo i
- R_i = suma de rangos del grupo i

6.5.3 Aplicación en producción animal

Caso Zootécnico: Comparación de tres dietas en cerdos de engorde

Se evalúa el efecto de tres dietas sobre la ganancia diaria de peso (GDP) en gramos durante 30 días en 18 cerdos.

Tabla 39. GDP según tipo de dieta.

Dieta A (n=6)	Dieta B (n=6)	Dieta C (n=6)
850, 880	920, 935	780, 795
865, 875	910, 925	770, 785
870, 890	915, 930	775, 790

Procedimiento de análisis:

1. **Asignación de rangos globales**
2. **Cálculo de sumas de rangos:**
 - $RA = 57$
 - $RB = 93$
 - $RC = 21$

3. **Cálculo del estadístico H:** $H = [12/(18 \times 19)] \times [(57^2/6) + (93^2/6) + (21^2/6)] - 3(19)$ $H = 14.82$
4. **Valor crítico:** $\chi^2(2, 0.05) = 5.99$
5. **Decisión:** $H = 14.82 > 5.99$, se rechaza H_0

Conclusión: Existen diferencias significativas ($p < 0.001$) entre las tres dietas en términos de ganancia diaria de peso.

6.5.4 Comparaciones Múltiples Post-hoc

Cuando se rechaza la hipótesis nula en Kruskal-Wallis, se requieren pruebas post-hoc para identificar qué grupos difieren. El método de Dunn con corrección de Bonferroni es ampliamente utilizado (Berlanga & Rubio Hurtado, 2012):

```
# Instalación del paquete necesario
install.packages("dunn.test")
library(dunn.test)

# Datos
gdp <- c(850,880,865,875,870,890,920,935,910,925,915,930,780,795,770,785,
775,790)
dieta <- factor(rep(c("A","B","C"), each=6))

# Prueba de Kruskal-Wallis
kruskal.test(gdp ~ dieta)

# Comparaciones post-hoc
dunn.test(gdp, dieta, method="bonferroni")
```

Figura 62. Implementación en R.

6.6 Prueba de Friedman

6.6.1 Fundamentos teóricos

La prueba de Friedman constituye la alternativa no paramétrica al ANOVA de medidas repetidas o al ANOVA de bloques completos al azar. Se utiliza cuando se tienen k tratamientos aplicados a los mismos sujetos o bloques, y los datos no cumplen los supuestos paramétricos (Ortiz, 2020).

6.6.2 Diseño y aplicación

El diseño típico involucra:

- b bloques (sujetos o unidades experimentales)
- k tratamientos aplicados a cada bloque
- Variable respuesta medida en escala ordinal o continua

El estadístico de Friedman se calcula como:

$$\chi^2_r = [12/bk(k+1)] \times \sum R_j^2 - 3b(k+1)$$

Donde R_j es la suma de rangos del tratamiento j.

6.6.3 Caso veterinario: Evaluación de analgésicos

Estudio del dolor postoperatorio en gatos

Se evalúan tres protocolos analgésicos en 8 gatos sometidos a ovariectomía, midiendo el dolor mediante escala visual análoga (0-10) a las 2, 6 y 12 horas post-cirugía.

Tabla 40. Scores de dolor según protocolo analgésico.

Gato	Protocolo A	Protocolo B	Protocolo C	Rangos A	Rangos B	Rangos C
1	6	3	4	3	1	2
2	7	4	5	3	1	2
3	5	2	3	3	1	2
4	6	3	4	3	1	2
5	8	4	5	3	1	2
6	7	3	4	3	1	2
7	6	2	3	3	1	2
8	7	4	5	3	1	2

Análisis:

- Sumas de rangos: RA = 24, RB = 8, RC = 16
- $\chi^2 = [12/(8 \times 3 \times 4)] \times (24^2 + 8^2 + 16^2) - 3 \times 8 \times 4 = 16$
- Grados de libertad = k - 1 = 2
- Valor crítico $\chi^2(2, 0.05) = 5.99$
- Decisión: $\chi^2 = 16 > 5.99$, se rechaza H₀

Interpretación clínica: Los tres protocolos analgésicos producen niveles de analgesia significativamente diferentes ($p < 0.001$) en gatos post-ovariohisterectomía.

6.7 Pruebas de Chi-cuadrado

6.7.1 Fundamentos y aplicaciones

La prueba de chi-cuadrado (χ^2) evalúa la asociación entre variables categóricas, siendo fundamental en estudios epidemiológicos veterinarios (Hernández de la Rosa et al., 2017). Se distinguen principalmente:

1. **Prueba de independencia:** Evalúa asociación entre dos variables categóricas
2. **Prueba de bondad de ajuste:** Compara frecuencias observadas con esperadas
3. **Prueba de homogeneidad:** Compara proporciones entre poblaciones

6.7.2 Cálculo del estadístico

$$\chi^2 = \sum [(O_i - E_i)^2 / E_i]$$

Donde:

- O_i = frecuencia observada en la celda i
- E_i = frecuencia esperada en la celda i

6.7.3 Aplicación epidemiológica

Estudio de caso: Asociación entre vacunación y enfermedad respiratoria en bovinos

Se investiga la asociación entre el estado de vacunación y la presencia de enfermedad respiratoria bovina en 200 animales.

Tabla 41. Tabla de contingencia 2×2.

Estado	Enfermo	Sano	Total
Vacunado	15	85	100
No vacunado	45	55	100
Total	60	140	200

Cálculo de frecuencias esperadas:

- $E_{11} = (100 \times 60)/200 = 30$
- $E_{12} = (100 \times 140)/200 = 70$
- $E_{21} = (100 \times 60)/200 = 30$
- $E_{22} = (100 \times 140)/200 = 70$

Estadístico chi-cuadrado: $\chi^2 = [(15-30)^2/30] + [(85-70)^2/70] + [(45-30)^2/30] + [(55-70)^2/70]$ $\chi^2 = 7.5 + 3.21 + 7.5 + 3.21 = 21.42$

Medidas de asociación:

1. **Riesgo Relativo (RR):** $RR = (15/100)/(45/100) = 0.33$
2. **Odds Ratio (OR):** $OR = (15 \times 55)/(85 \times 45) = 0.216$
3. **Reducción del Riesgo Absoluto:** $RRA = 45\% - 15\% = 30\%$

Interpretación epidemiológica: Existe una asociación significativa ($\chi^2 = 21.42$, $p < 0.001$) entre la vacunación y la protección contra

enfermedad respiratoria. Los animales vacunados tienen 67% menos riesgo de enfermedad (RR = 0.33).

6.7.4 Corrección de Yates

Para tablas 2×2 con frecuencias esperadas < 5, se aplica la corrección de continuidad de Yates:

$$\chi^2_{\text{Yates}} = \sum [(|O_i - E_i| - 0.5)^2 / E_i]$$

6.8 Prueba exacta de Fisher

6.8.1 Fundamentos teóricos

La prueba exacta de Fisher proporciona una alternativa a la prueba chi-cuadrado cuando los tamaños muestrales son pequeños o cuando alguna frecuencia esperada es menor a 5 (Molina, 2021). Esta prueba calcula la probabilidad exacta de obtener los resultados observados bajo la hipótesis nula de independencia.

6.8.2 Aplicación en investigación clínica

Caso: Evaluación de tratamiento para dermatitis atópica canina

Se evalúa un nuevo tratamiento inmunomodulador en 12 perros con dermatitis atópica severa.

Tabla 42. Respuesta al tratamiento.

Tratamiento	Mejoría	Sin mejoría	Total
-------------	---------	-------------	-------

Nuevo	6	1	7
Convencional	1	4	5
Total	7	5	12

Cálculo de probabilidad exacta:

$$P = [(a+b)!(c+d)!(a+c)!(b+d)!] / [n!a!b!c!d!]$$

$$P = [(7! \times 5! \times 7! \times 5!)] / [12! \times 6! \times 1! \times 1! \times 4!] \quad P = 0.015$$

Interpretación: La probabilidad exacta ($p = 0.015$) indica diferencias significativas entre tratamientos, sugiriendo superioridad del nuevo protocolo inmunomodulador.

6.8.3 Implementación computacional

```
# Datos del ejemplo
tabla <- matrix(c(6,1,1,4), nrow=2,
  dimnames=list(Tratamiento=c("Nuevo","Convencional"),
    Respuesta=c("Mejoria","Sin_mejoria")))

# Prueba exacta de Fisher
fisher.test(tabla)

# Visualización
library(ggplot2)
library(reshape2)
datos_long <- melt(tabla)
ggplot(datos_long, aes(x=Respuesta, y=value, fill=Tratamiento)) +
  geom_bar(stat="identity", position="dodge") +
  labs(y="Número de casos", title="Respuesta al tratamiento de dermatitis")
```

Figura 63. Implementación computacional.

6.9 Correlación de Spearman

6.9.1 Fundamentos teóricos

El coeficiente de correlación de Spearman (ρ o r_s) constituye una medida no paramétrica de asociación monotónica entre dos variables, basada en los rangos de las observaciones en lugar de sus valores originales (Mendivelso, 2021). Esta característica lo hace robusto ante valores atípicos y aplicable a datos ordinales.

6.9.2 Cálculo y propiedades

El coeficiente de Spearman se calcula mediante:

$$\rho = 1 - [6\sum d_i^2]/[n(n^2-1)]$$

Donde:

- d_i = diferencia entre rangos de las observaciones pareadas
- n = número de pares

Propiedades:

- Rango: $-1 \leq \rho \leq 1$
- $\rho = 1$: correlación monotónica positiva perfecta
- $\rho = -1$: correlación monotónica negativa perfecta
- $\rho = 0$: ausencia de correlación monotónica

6.9.3 Aplicación en medicina veterinaria

Estudio: Relación entre condición corporal y parámetros reproductivos en vacas lecheras

Se evalúa la asociación entre el score de condición corporal (SCC, escala 1-5) y el intervalo parto-concepción (días) en 15 vacas Holstein.

Tabla 43. Datos de SCC e intervalo parto-concepción.

Vaca	SCC	Días	Rango SCC	Rango Días	di	di ²
1	2.5	125	4	11	-7	49
2	3.0	95	8	5	3	9
3	2.0	145	1	14	-13	169
4	3.5	85	11	3	8	64
5	2.5	120	5	10	-5	25
6	4.0	75	14	1	13	169
7	2.25	135	2	13	-11	121
8	3.25	90	9.5	4	5.5	30.25
9	2.75	110	6	8	-2	4
10	3.75	80	12.5	2	10.5	110.25
11	3.0	100	7	6	1	1
12	2.25	130	3	12	-9	81

13	3.25	105	9.5	7	2.5	6.25
14	3.75	115	12.5	9	3.5	12.25
15	4.25	150	15	15	0	0

Cálculo: $\Sigma di^2 = 851.5$ $\rho = 1 - [6(851.5)]/[15(225-1)] = 1 - 5109/3360 = -0.52$

Prueba de significancia: $t = \rho\sqrt{[(n-2)/(1-\rho^2)]} = -0.52\sqrt{[13/0.73]} = -2.19$
Grados de libertad = 13 Valor crítico ($\alpha = 0.05$, bilateral) = ± 2.16

Interpretación zootécnica: Existe una correlación negativa moderada y significativa ($\rho = -0.52$, $p < 0.05$) entre la condición corporal y el intervalo parto-concepción, indicando que vacas con mejor condición corporal presentan intervalos más cortos para quedar gestantes nuevamente.

6.9.4 Comparación con correlación de Pearson

Ventajas de Spearman:

- No requiere normalidad bivariada
- Robusto ante valores atípicos
- Aplicable a datos ordinales
- Detecta relaciones monotónicas no lineales

```

# Generación de datos con relación no lineal
set.seed(123)
peso <- seq(20, 80, length.out = 30)
glucosa <- 70 + 2*peso + 0.05*peso^2 + rnorm(30, 0, 10)

# Correlaciones
cor_pearson <- cor(peso, glucosa, method = "pearson")
cor_spearman <- cor(peso, glucosa, method = "spearman")

# Visualización
par(mfrow = c(1, 2))
plot(peso, glucosa, main = paste("Pearson r =", round(cor_pearson,
3)),
      xlab = "Peso (kg)", ylab = "Glucosa (mg/dL)")
abline(lm(glucosa ~ peso), col = "red")

plot(rank(peso), rank(glucosa), main = paste("Spearman  $\rho$  =",
round(cor_spearman, 3)),
      xlab = "Rango Peso", ylab = "Rango Glucosa")
abline(lm(rank(glucosa) ~ rank(peso)), col = "blue")

```

Figura 64. Ejemplo comparativo con software.

6.10 Actividades de autoevaluación

6.10.1 Ejercicios resueltos

Ejercicio 1: Selección de prueba apropiada

Un investigador desea comparar la eficacia de tres vermífugos en equinos. Cuenta con 18 caballos divididos aleatoriamente en tres grupos. La variable respuesta es el porcentaje de reducción en el conteo de huevos de parásitos, que presenta distribución asimétrica.

Solución:

- Variable: continua con distribución no normal
- Diseño: tres grupos independientes
- Prueba apropiada: Kruskal-Wallis
- Justificación: Alternativa no paramétrica al ANOVA de una vía cuando no se cumple normalidad

Ejercicio 2: Interpretación de resultados

En un estudio sobre mastitis bovina, se obtiene $\chi^2 = 8.45$ con 1 grado de libertad al evaluar la asociación entre sistema de ordeño (manual vs. mecánico) y presencia de mastitis.

Solución:

- Valor crítico $\chi^2(1, 0.05) = 3.84$
- Como $8.45 > 3.84$, $p < 0.05$
- Conclusión: Existe asociación significativa entre el sistema de ordeño y la presencia de mastitis

Problemas Propuestos

1. **Nivel básico:** Un veterinario registra el tiempo de recuperación (minutos) post-anestesia en 10 gatos: 5 machos (25, 30, 28, 32, 27) y 5 hembras (35, 38, 40, 36, 41). Aplique la prueba de Mann-Whitney para evaluar diferencias entre sexos.
2. **Nivel intermedio:** Se evalúan tres suplementos nutricionales en la ganancia de peso de conejos. Diseñe un experimento con 12 animales divididos en 4 bloques según peso inicial, aplicando

cada suplemento a cada bloque. Indique qué prueba no paramétrica utilizaría y justifique.

3. **Nivel avanzado:** En un estudio longitudinal, se mide la concentración de cortisol (nmol/L) en 20 perros de refugio durante tres momentos: ingreso, semana 1 y semana 4. Los datos muestran fuerte asimetría positiva. Proponga el análisis estadístico completo incluyendo comparaciones post-hoc si corresponde.

6.11 Guía de Implementación en Software estadístico

6.11.1 Aplicación Integral con R

Esta sección proporciona una guía completa para implementar todas las pruebas no paramétricas del capítulo utilizando R, con énfasis en la interpretación veterinaria de los resultados.

```

# Instalación y carga de paquetes necesarios
packages <- c("tidyverse", "dunn.test", "FSA", "rstatix",
              "ggpubr", "corrplot", "pwr", "effectsize")

# Función para instalar si no están disponibles
install_if_missing <- function(packages) {
  new_packages <- packages[!(packages %in%
installed.packages()[,"Package"])]
  if(length(new_packages)) install.packages(new_packages)
  lapply(packages, library, character.only = TRUE)
}

install_if_missing(packages)

# Configuración de opciones
options(scipen = 999) # Evitar notación científica
set.seed(2024) # Reproducibilidad

```

Figura 65. Configuración del entorno de trabajo.

6.11.2 Análisis completo: Caso de estudio multivariado

Escenario clínico: Evaluación de tres protocolos anestésicos en cirugía de tejidos blandos en caninos.

```

# Simulación de datos basados en caso real
n_por_grupo <- 20
protocolos <- factor(rep(c("Propofol", "Ketamina", "Alfaxalona"),
each = n_por_grupo))

# Variables medidas
tiempo_induccion <- c(
  rnorm(20, 45, 8), # Propofol
  rnorm(20, 65, 12), # Ketamina
  rnorm(20, 50, 10) # Alfaxalona
)

# Score de calidad de recuperación (1-10)
calidad_recuperacion <- c(
  sample(6:9, 20, replace = TRUE), # Propofol
  sample(4:7, 20, replace = TRUE), # Ketamina
  sample(7:10, 20, replace = TRUE) # Alfaxalona
)

# Crear dataframe
datos_anestesia <- data.frame(
  paciente_id = 1:60,
  protocolo = protocolos,
  tiempo_induccion = round(tiempo_induccion, 1),
  calidad_recuperacion = calidad_recuperacion,
  peso = round(rnorm(60, 25, 8), 1),
  edad_meses = sample(12:120, 60, replace = TRUE)
)

# Exploración inicial
summary(datos_anestesia)

```

Figura 66. Caso de estudio multivariado.

6.11.3 Análisis de normalidad y decisión de prueba

```
# Función para evaluación completa de normalidad
evaluar_normalidad <- function(data, variable, grupo) {
  # Prueba de Shapiro-Wilk por grupo
  shapiro_results <- data %>%
    group_by({{grupo}}) %>%
    summarise(
      n = n(),
      W_statistic = shapiro.test({{variable}})$statistic,
      p_value = shapiro.test({{variable}})$p.value,
      normal = ifelse(p_value > 0.05, "Sí", "No")
    )

  # Visualización
  p1 <- ggplot(data, aes(x = {{variable}}, fill = {{grupo}})) +
    geom_histogram(alpha = 0.6, position = "identity", bins = 15) +
    facet_wrap(~{{grupo}}, scales = "free_y") +
    theme_minimal() +
    labs(title = "Distribución por grupo")

  p2 <- ggplot(data, aes(sample = {{variable}})) +
    stat_qq() +
    stat_qq_line() +
    facet_wrap(~{{grupo}}) +
    theme_minimal() +
    labs(title = "Q-Q Plot")

  return(list(
    tabla = shapiro_results,
    histograma = p1,
    qqplot = p2
  ))
}

# Aplicar evaluación
normalidad_tiempo <- evaluar_normalidad(datos_anestesia,
tiempo_induccion, protocolo)
print(normalidad_tiempo$tabla)
```

Figura 67. Análisis de normalidad y decisión de prueba.

6.11.4 Aplicación de Kruskal-Wallis con Post-hoc

```
# Análisis principal
kw_resultado <- kruskal.test(tiempo_induccion ~ protocolo, data = datos_anestesia)
# Función mejorada para reporte
reportar_kruskal_wallis <- function(resultado, data, variable, grupo) {
  cat("=" * 50, "\n")
  cat("ANÁLISIS DE KRUSKAL-WALLIS\n")
  cat("=" * 50, "\n\n")
  # Estadísticas descriptivas
  desc_stats <- data %>%
    group_by({grupo}) %>%
    summarise(
      n = n(),
      mediana = median({variable}),
      Q1 = quantile({variable}, 0.25),
      Q3 = quantile({variable}, 0.75),
      rango_medio = mean(rank({variable}))
    )
  print(desc_stats)
  cat("\nEstadístico H: ", round(resultado$statistic, 3), "\n")
  cat("Grados de libertad: ", resultado$parameter, "\n")
  cat("Valor p: ", format(resultado$p.value, scientific = FALSE), "\n")
  # Tamaño del efecto (Epsilon-squared)
  n_total <- nrow(data)
  epsilon_sq <- (resultado$statistic - resultado$parameter + 1) / (n_total -
    resultado$parameter)
  cat("Epsilon-squared: ", round(epsilon_sq, 3), "\n")
  cat("Interpretación del tamaño del efecto: ")
  if(epsilon_sq < 0.01) cat("Negligible\n")
  else if(epsilon_sq < 0.06) cat("Pequeño\n")
  else if(epsilon_sq < 0.14) cat("Moderado\n")
  else cat("Grande\n")
  # Decisión estadística
  cat("\nDecisión: ") if(resultado$p.value < 0.05) {
    cat("Se rechaza H0. Existen diferencias significativas entre grupos.\n")
  } else {
    cat("No se rechaza H0. No hay evidencia de diferencias entre grupos.\n")
  }
  reportar_kruskal_wallis(kw_resultado, datos_anestesia, tiempo_induccion, protocolo)
# Comparaciones post-hoc si es significativo
if(kw_resultado$p.value < 0.05) {
  cat("\n", "=" * 50, "\n")
  cat("COMPARACIONES POST-HOC (Método de Dunn)\n")
  cat("=" * 50, "\n\n")
  dunn_result <- dunn.test(datos_anestesia$tiempo_induccion,
    datos_anestesia$protocolo,
    method = "bonferroni")
  # Crear tabla de resultados formateada
  comparaciones <- data.frame(
    Comparación = dunn_result$comparisons,
    Z = round(dunn_result$Z, 3),
    P_ajustado = round(dunn_result$P.adjusted, 4),
    Significativo = ifelse(dunn_result$P.adjusted < 0.05, "Sí", "No"))
  print(comparaciones)}
```

Figura 68. Aplicación de Kruskal-Wallis con Post-hoc.

6.11.5 Visualización profesional de resultados

```
# Función para crear gráficos de publicación
crear_boxplot_veterinario <- function(data, x, y, titulo, y_label) {
  # Calcular significancias
  stat_test <- data %>%
    rstatix::kruskal_test(formula = as.formula(paste(deparse(substitute(y)), "~",
    deparse(substitute(x)))))
  # Comparaciones pareadas
  pwc <- data %>%
    rstatix::dunn_test(formula = as.formula(paste(deparse(substitute(y)), "~",
    deparse(substitute(x)))),
    p.adjust.method = "bonferroni")
  pwc <- pwc %>% add_xy_position(x = deparse(substitute(x)))
  # Crear gráfico
  p <- ggplot(data, aes(x = {{x}}, y = {{y}}, fill = {{x}})) +
    geom_boxplot(alpha = 0.7, outlier.shape = NA) +
    geom_jitter(width = 0.2, alpha = 0.5, size = 2) +
    stat_summary(fun = mean, geom = "point", shape = 18, size = 4, color = "red") +
    scale_fill_brewer(palette = "Set2") +
    labs(title = titulo,
    subtitle = paste("Kruskal-Wallis: H =", round(stat_test$statistic, 2),
    ", p =", format(stat_test$p, scientific = FALSE)),
    y = y_label, x = "Protocolo Anestésico")
  p + theme_classic() + theme(plot.title = element_text(size = 14, face = "bold"),
  axis.text = element_text(size = 11),
  axis.title = element_text(size = 12),
  legend.position = "none")
}
# Agregar significancias si existen
if(any(pwc$p.adj < 0.05)) {
  p <- p + stat_pvalue_manual(
    pwc[pwc$p.adj < 0.05,],
    label = "p.adj.signif",
    tip.length = 0.01,
    step.increase = 0.08
  )
}
return(p)
}
# Aplicar función
plot_tiempo <- crear_boxplot_veterinario(
  datos_anestesia,
  protocolo,
  tiempo_induccion,
  "Comparación de Tiempo de Inducción Anestésica",
  "Tiempo de Inducción (segundos)")
print(plot_tiempo)
# Guardar gráfico en alta calidad
ggsave("tiempo_induccion_anestesia.pdf", plot_tiempo,
  width = 10, height = 6, dpi = 300)
```

Figura 69. Visualización profesional de resultados.

6.12 Conclusiones

El presente capítulo ha abordado de manera comprehensiva las pruebas no paramétricas más relevantes en la investigación veterinaria, proporcionando al estudiante y al profesional las herramientas estadísticas necesarias para el análisis de datos que no cumplen los supuestos de normalidad o que presentan escalas de medición ordinales. La aplicación de estas técnicas en medicina veterinaria y zootecnia resulta particularmente valiosa considerando las características frecuentes de los datos clínicos y experimentales en estas disciplinas, donde los tamaños muestrales suelen ser limitados por consideraciones éticas, económicas y logísticas.

La prueba de Mann-Whitney U se ha establecido como una alternativa robusta y confiable a la prueba t de Student para comparar dos grupos independientes. Su aplicación en casos como la evaluación de tratamientos antiparasitarios o la comparación de protocolos terapéuticos demuestra su versatilidad y relevancia práctica. De manera similar, la prueba de Wilcoxon para muestras pareadas ofrece una herramienta invaluable para diseños antes-después, comunes en la evaluación de intervenciones clínicas donde el mismo animal sirve como su propio control, minimizando la variabilidad interindividual y optimizando el uso de recursos experimentales.

Las pruebas de Kruskal-Wallis y Friedman extienden estas capacidades analíticas a diseños más complejos con múltiples grupos o medidas repetidas, respectivamente. Su importancia en estudios de nutrición animal, evaluación de protocolos anestésicos y ensayos farmacológicos

veterinarios no puede ser subestimada. La inclusión de procedimientos post-hoc, particularmente la prueba de Dunn con corrección de Bonferroni, permite identificar diferencias específicas entre grupos manteniendo el control del error tipo I, aspecto crucial para la validez de las conclusiones en investigación biomédica.

El análisis de variables categóricas mediante las pruebas de chi-cuadrado y la prueba exacta de Fisher proporciona herramientas fundamentales para estudios epidemiológicos y de medicina preventiva veterinaria. La capacidad de evaluar asociaciones entre factores de riesgo y presencia de enfermedad, calcular medidas de asociación como el riesgo relativo y el odds ratio, y determinar la eficacia de intervenciones preventivas como programas de vacunación, constituye una competencia esencial para el veterinario moderno orientado hacia la medicina basada en evidencia.

La correlación de Spearman complementa este conjunto de herramientas al permitir evaluar asociaciones monotónicas entre variables continuas u ordinales sin asumir linealidad o normalidad bivariada. Su robustez ante valores atípicos y su aplicabilidad a escalas de medición ordinales, frecuentes en la evaluación clínica veterinaria mediante scores y escalas visuales, la convierten en una técnica indispensable para el análisis de relaciones entre variables en contextos clínicos y productivos.

La implementación práctica de estas técnicas mediante software estadístico especializado representa un avance significativo en la democratización del análisis estadístico riguroso en ciencias

veterinarias. La disponibilidad de plataformas como R, Python y SPSS, junto con las guías de implementación proporcionadas en este capítulo, permite que tanto estudiantes como profesionales puedan realizar análisis sofisticados sin requerir conocimientos matemáticos avanzados, aunque manteniendo la comprensión conceptual necesaria para una interpretación apropiada de los resultados.

Es fundamental enfatizar que la elección entre pruebas paramétricas y no paramétricas no debe basarse únicamente en el cumplimiento o incumplimiento de supuestos estadísticos, sino también en consideraciones sobre el diseño experimental, la naturaleza de las variables medidas y los objetivos específicos de la investigación. Las pruebas no paramétricas, aunque generalmente menos poderosas que sus contrapartes paramétricas cuando se cumplen todos los supuestos, ofrecen mayor robustez y confiabilidad en situaciones comunes en la práctica veterinaria donde dichos supuestos son cuestionables.

La integración de consideraciones éticas en el análisis estadístico, particularmente en lo referente al cálculo apropiado del tamaño muestral y el reporte transparente de resultados, refleja la evolución hacia una investigación veterinaria más responsable y rigurosa. El principio de las 3Rs (Reemplazo, Reducción y Refinamiento) en experimentación animal se ve apoyado por el uso apropiado de técnicas estadísticas que maximizan la información obtenida de cada animal utilizado en investigación, minimizando así el número total de animales requeridos mientras se mantiene el poder estadístico adecuado.

Las aplicaciones prácticas presentadas a lo largo del capítulo, desde la evaluación de tratamientos en pequeños animales hasta estudios de producción en especies mayores, demuestran la universalidad de estas técnicas en el amplio espectro de las ciencias veterinarias. La capacidad de analizar datos desde ensayos clínicos controlados hasta estudios observacionales epidemiológicos, pasando por experimentos de nutrición y producción animal, subraya la importancia de dominar estas herramientas para el ejercicio profesional integral.

CAPÍTULO VII

7 ANÁLISIS DE CORRELACIÓN Y REGRESIÓN

7.1 Objetivos del Capítulo

Al finalizar este capítulo, el estudiante será capaz de:

1. Comprender los conceptos fundamentales de correlación y regresión en el contexto veterinario
2. Calcular e interpretar coeficientes de correlación de Pearson en variables clínicas y productivas
3. Desarrollar modelos de regresión lineal simple y múltiple para predicción de parámetros veterinarios
4. Aplicar regresión logística en estudios epidemiológicos y diagnósticos clínicos
5. Realizar diagnósticos completos de modelos estadísticos y validar sus supuestos
6. Implementar análisis de correlación y regresión utilizando software estadístico especializado

7.2 Introducción

El análisis de correlación y regresión constituye una herramienta fundamental en la investigación veterinaria moderna. Estas técnicas estadísticas permiten explorar relaciones entre variables, predecir resultados clínicos y optimizar decisiones terapéuticas y productivas. En la práctica veterinaria, la capacidad de establecer relaciones cuantitativas entre variables como peso corporal y dosis farmacológicas,

edad y parámetros fisiológicos, o factores ambientales y producción animal, resulta esencial para la toma de decisiones basadas en evidencia.

La correlación examina la fuerza y dirección de la relación entre dos variables continuas, mientras que la regresión permite modelar esta relación para realizar predicciones. Ambas técnicas se complementan y proporcionan información valiosa sobre los fenómenos biológicos que caracterizan la medicina veterinaria y la producción animal.

7.3 Correlación de Pearson

7.3.1 Fundamentos teóricos

El coeficiente de correlación de Pearson (r) cuantifica la relación lineal entre dos variables continuas. Este estadístico, desarrollado por Karl Pearson, varía entre -1 y +1, donde valores cercanos a +1 indican correlación positiva perfecta, valores cercanos a -1 señalan correlación negativa perfecta, y valores próximos a 0 sugieren ausencia de relación lineal (Apaza Zúñiga et al., 2022).

La fórmula del coeficiente de correlación de Pearson se expresa como:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Donde:

- x_i e y_i representan los valores individuales de las variables X e Y
- $\bar{x}$ e $\bar{y}$ son las medias de X e Y respectivamente

- n es el tamaño de la muestra

7.3.2 Supuestos y condiciones de aplicación

Para que el coeficiente de correlación de Pearson proporcione resultados válidos, deben cumplirse los siguientes supuestos (Peñaherrera-Larenas et al., 2021):

1. **Linealidad:** La relación entre las variables debe ser aproximadamente lineal
2. **Normalidad bivariada:** Ambas variables deben seguir una distribución normal conjunta
3. **Homocedasticidad:** La varianza debe ser constante a lo largo del rango de valores
4. **Independencia de observaciones:** Cada par de datos debe ser independiente
5. **Variables continuas:** Ambas variables deben medirse en escala de intervalo o razón

7.3.3 Interpretación del coeficiente de correlación

La interpretación del coeficiente de correlación en investigación veterinaria sigue criterios establecidos, aunque debe contextualizarse según el fenómeno estudiado:

Tabla 44. Interpretación del Coeficiente de Correlación de Pearson.

Valor de r	Interpretación	Ejemplo Veterinario
--------------	----------------	---------------------

0.90 1.00	a	Correlación muy alta positiva	Peso corporal vs. dosis de anestésico en caninos
0.70 0.89	a	Correlación alta positiva	Consumo de alimento vs. ganancia de peso en bovinos
0.50 0.69	a	Correlación moderada positiva	Edad vs. niveles de creatinina en felinos geriátricos
0.30 0.49	a	Correlación baja positiva	Temperatura ambiental vs. consumo de agua en equinos
-0.29 0.29	a	Correlación despreciable	Color del pelaje vs. temperamento en caninos
-0.30 0.49	a -	Correlación baja negativa	Ejercicio vs. glucemia en perros diabéticos
-0.50 0.69	a -	Correlación moderada negativa	Densidad poblacional vs. peso al destete en porcinos
-0.70 0.89	a -	Correlación alta negativa	Edad materna vs. tamaño de camada en conejos
-0.90 1.00	a -	Correlación muy alta negativa	Parasitemia vs. hematocrito en babesiosis canina

7.3.4 Ejemplo práctico: Correlación entre peso corporal y frecuencia cardíaca en caninos

Caso de Estudio: Un hospital veterinario desea investigar la relación entre el peso corporal y la frecuencia cardíaca en reposo en perros adultos sanos.

Paciente	Peso (kg)	FC (lpm)
1	5.2	140
2	8.7	125
3	12.3	115
4	15.8	108
5	20.4	95
...	...	...

Figura 70. Datos recolectados (n=20).

Análisis paso a paso:

1. Cálculo de medias:

- $\bar{x}$ peso=18.5 kg
- $\bar{y}_{FC}$ =102.3l pm

2. Cálculo de desviaciones y productos:

- $\sum(x_i - \bar{x})(y_i - \bar{y}) = -2845.6$
- $\sum(x_i - \bar{x})^2 = 892.4$
- $\sum(y_i - \bar{y})^2 = 3256.8$

3. Aplicación de la fórmula: $r = \frac{-2845.6}{\sqrt{892.4 \times 3256.8}} = -0.847$

Interpretación: Existe una correlación alta negativa ($r = -0.847$) entre el peso corporal y la frecuencia cardíaca en caninos, indicando que perros de mayor peso tienden a presentar frecuencias cardíacas más bajas en reposo, consistente con los principios fisiológicos de la alometría cardiovascular.

7.3.5 Prueba de significancia

Para determinar si la correlación observada es estadísticamente significativa, se emplea la prueba t de Student:

$$t = r \sqrt{\frac{n-2}{1-r^2}}$$

Con n-2 grados de libertad.

Para el ejemplo anterior:

$$t = -0.847 \sqrt{\frac{18}{1-0.717}} = -6.73$$

Con 18 grados de libertad y $\alpha = 0.05$, el valor crítico es ± 2.101 . Como $|t| > t$ crítico, se rechaza la hipótesis nula de no correlación ($p < 0.001$).

7.4 Regresión lineal simple

7.4.1 Modelo de regresión lineal simple

La regresión lineal simple modela la relación entre una variable independiente (X) y una variable dependiente (Y) mediante la ecuación:

$$Y = \beta_0 + \beta_1 X + \epsilon$$

Donde:

- Y = Variable dependiente o respuesta
- X = Variable independiente o predictora

- β_0 = Intercepto (valor de Y cuando X = 0)
- β_1 = Pendiente (cambio en Y por unidad de cambio en X)
- ε = Error aleatorio

7.4.2 Estimación de parámetros por mínimos cuadrados

Los estimadores de mínimos cuadrados se calculan mediante:

$$r = \frac{\sum_{i=1}^n (xi - \bar{x})(yi - \bar{y})}{\sum_{i=1}^n (xi - \bar{x})^2}$$

$$\widehat{\beta}_0 = \bar{y} - \widehat{\beta}_1 \bar{x}$$

7.4.3 Ejemplo aplicado: Predicción de glucemia según peso en felinos diabéticos

Contexto Clínico: Un veterinario especialista en endocrinología felina desea desarrollar un modelo predictivo para estimar los niveles de glucemia en ayunas basándose en el peso corporal de gatos diabéticos.

Datos del estudio (n=30 gatos diabéticos):

Tabla 45. Datos de Peso y Glucemia en Felinos Diabéticos

Peso (kg)	Glucemia (mg/dL)
3.8	285
4.2	298
4.5	312

5.1	325
5.6	338
...	...

Desarrollo del modelo:

1. Cálculo de estadísticos descriptivos:

- Media peso: $\bar{x}=5.2$ kg
- Media glucemia: $\bar{y}=320$ mg/dL
- Desviación estándar peso: $s_x=1.1$ kg
- Desviación estándar glucemia: $s_y=42$ mg/dL

2. Estimación de parámetros:

$$\widehat{\beta}_1 = \frac{1458.6}{36.3} = 40.2$$

$$\widehat{\beta}_0 = 320 - (40.2 \times 5.2) = 111.0$$

3. Ecuación de regresión: $\text{Glucemia} = 111.0 + 40.2 \times \text{Peso}$

Interpretación clínica: Por cada kilogramo adicional de peso corporal, la glucemia en ayunas aumenta en promedio 40.2 mg/dL en gatos diabéticos. Este modelo permite al veterinario anticipar niveles glucémicos y ajustar protocolos de insulino terapia.

7.4.4 Coeficiente de determinación (R^2)

El coeficiente de determinación indica la proporción de variabilidad en Y explicada por el modelo:

$$R = \frac{SSR}{SST} = \frac{\sum(y_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}$$

Para el ejemplo de felinos diabéticos: $R^2 = 0.724$, indicando que el 72.4% de la variabilidad en los niveles de glucemia se explica por el peso corporal.

7.4.5 Intervalos de confianza y predicción

Intervalo de confianza para la respuesta media:

$$\hat{y} \pm t_{\frac{\alpha}{2}, n-2} x s_e \sqrt{\frac{1}{n} + \frac{(x_o - \bar{x})^2}{\sum(x_i - \bar{x})^2}}$$

Intervalo de predicción para una observación individual:

$$\hat{y} \pm t_{\frac{\alpha}{2}, n-2} x s_e \sqrt{1 + \frac{1}{n} + \frac{(x_o - \bar{x})^2}{\sum(x_i - \bar{x})^2}}$$

7.5 Regresión lineal múltiple

7.5.1 Modelo general

La regresión lineal múltiple extiende el modelo simple incorporando múltiples variables predictoras:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k \in$$

En notación matricial:

$$Y = X\beta + \varepsilon$$

7.5.2 Estimación de parámetros

Los estimadores de mínimos cuadrados se obtienen mediante:

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

7.5.3 Caso práctico: Predicción de producción láctea en bovinos

Objetivo: Desarrollar un modelo predictivo para la producción láctea diaria considerando múltiples factores.

Variables del estudio:

- Y: Producción láctea (litros/día)
- X₁: Peso corporal (kg)
- X₂: Edad (meses)
- X₃: Consumo de materia seca (kg/día)
- X₄: Número de lactancia
- X₅: Días en lactación

Tabla 46. Análisis de varianza del modelo de regresión múltiple.

Fuente	GL	SC	CM	F	Valor p
Regresión	5	2856.4	571.3	45.2	<0.001
Error	144	1821.6	12.6		
Total	149	4678.0			

R^2 ajustado = 0.598

7.5.4 Interpretación de coeficientes

Los coeficientes de regresión parcial representan el cambio en Y por unidad de cambio en X_i , manteniendo constantes las demás variables:

- **$\beta_1 = 0.042$:** Por cada kg adicional de peso, la producción aumenta 0.042 L/día
- **$\beta_2 = -0.15$:** Por cada mes adicional de edad, la producción disminuye 0.15 L/día
- **$\beta_3 = 1.85$:** Por cada kg adicional de materia seca consumida, la producción aumenta 1.85 L/día
- **$\beta_4 = 2.1$:** Cada lactancia adicional aumenta la producción en 2.1 L/día
- **$\beta_5 = -0.08$:** Por cada día adicional en lactación, la producción disminuye 0.08 L/día

7.5.5 Multicolinealidad

La multicolinealidad ocurre cuando las variables predictoras están altamente correlacionadas entre sí. Se detecta mediante (Oñate & Morales-Oñate, 2023):

Factor de Inflación de la Varianza (VIF):

$$VIP_j = \frac{1}{1 - R_j^2}$$

Donde R_j^2 es el coeficiente de determinación de la regresión de X_j sobre las demás variables independientes.

Criterios de interpretación:

- $VIF < 5$: No hay problemas de multicolinealidad
- $5 \leq VIF < 10$: Multicolinealidad moderada
- $VIF \geq 10$: Multicolinealidad severa

7.6 Regresión logística

7.6.1 Fundamentos del modelo logístico

La regresión logística modela la probabilidad de ocurrencia de un evento binario (Yupari-Azabache et al., 2021):

7.7 Conclusiones

El análisis de correlación y regresión constituye una piedra angular en la investigación y práctica veterinaria contemporánea, proporcionando herramientas estadísticas robustas que permiten transformar datos observacionales en conocimiento accionable. A lo largo de este capítulo, se ha desarrollado un marco comprehensivo que abarca desde los fundamentos de la correlación de Pearson hasta las complejidades de la regresión logística, siempre manteniendo un enfoque práctico y contextualizado en escenarios clínicos y productivos relevantes para el profesional veterinario. Esta aproximación integral no solo facilita la comprensión teórica de los conceptos, sino que prepara al estudiante para su aplicación inmediata en situaciones reales de la práctica profesional. La correlación, como medida fundamental de asociación entre variables, proporciona el punto de partida esencial para comprender las relaciones que gobiernan los fenómenos biológicos en medicina veterinaria. El dominio del coeficiente de correlación de Pearson permite al veterinario identificar y cuantificar patrones en datos clínicos, desde la relación entre peso corporal y parámetros fisiológicos hasta las asociaciones entre factores ambientales y productividad animal. Esta capacidad de reconocer y medir asociaciones constituye el primer paso hacia la construcción de modelos predictivos más complejos, estableciendo las bases para una práctica basada en evidencia cuantitativa. La comprensión profunda de los supuestos subyacentes y las limitaciones de la correlación asegura su aplicación apropiada y la interpretación correcta de los resultados, evitando las trampas comunes de confundir correlación con causalidad. Los modelos de regresión lineal, tanto simple como múltiple, representan la

evolución natural del análisis de correlación hacia la predicción y el modelado. La capacidad de desarrollar ecuaciones predictivas tiene implicaciones profundas en la práctica veterinaria, permitiendo desde el cálculo preciso de dosis farmacológicas hasta la predicción de resultados quirúrgicos y la optimización de parámetros productivos. La regresión lineal simple proporciona una herramienta elegante y poderosa para modelar relaciones bivariadas, mientras que la regresión múltiple abraza la complejidad multifactorial inherente a los sistemas biológicos. Esta progresión metodológica refleja la realidad de la medicina veterinaria, donde raramente un solo factor determina un resultado, y donde la consideración simultánea de múltiples variables es esencial para una comprensión completa del fenómeno estudiado. La introducción de la regresión logística amplía significativamente el alcance analítico del veterinario, permitiendo el modelado de resultados binarios tan comunes en la práctica clínica y la investigación epidemiológica. La capacidad de predecir probabilidades de eventos como supervivencia post-quirúrgica, respuesta a tratamiento o presencia de enfermedad, transforma la toma de decisiones clínicas de un proceso intuitivo a uno fundamentado en modelos estadísticos rigurosos. La interpretación de odds ratios proporciona una métrica intuitiva y clínicamente relevante del efecto de diferentes factores de riesgo, facilitando la comunicación de resultados tanto a colegas como a propietarios de animales. Esta técnica resulta particularmente valiosa en el contexto de la medicina veterinaria moderna, donde la estratificación del riesgo y la medicina personalizada ganan cada vez mayor importancia. El énfasis puesto en el diagnóstico y validación de modelos refleja un compromiso fundamental con el rigor científico y la

aplicabilidad práctica. El análisis de residuos, la evaluación de multicolinealidad, las pruebas de bondad de ajuste y la validación cruzada no son meros ejercicios académicos, sino salvaguardas esenciales que aseguran la confiabilidad y generalización de los modelos desarrollados. Esta aproximación crítica y reflexiva hacia el modelado estadístico cultiva en el estudiante una mentalidad científica rigurosa, donde cada conclusión está respaldada por evidencia sólida y cada predicción viene acompañada de una comprensión clara de su incertidumbre asociada. La capacidad de identificar y abordar violaciones de supuestos, valores atípicos influyentes y problemas de sobreajuste distingue al profesional competente del usuario casual de estadística. La integración de ejemplos prácticos específicos del ámbito veterinario a lo largo del capítulo cumple una función pedagógica crucial, demostrando la aplicabilidad inmediata de las técnicas presentadas. Desde la correlación entre peso y frecuencia cardíaca en caninos hasta modelos complejos de predicción de respuesta a terapias regenerativas en equinos, cada ejemplo ilustra no solo la mecánica del análisis sino también su interpretación clínica y sus implicaciones prácticas. Esta contextualización constante facilita la transferencia del conocimiento teórico hacia la aplicación práctica, preparando al estudiante para enfrentar los desafíos analíticos que encontrará en su carrera profesional. Los casos de estudio presentados abarcan diversas especies y contextos, desde pequeños animales en práctica clínica hasta sistemas de producción intensiva, reflejando la diversidad del campo veterinario. El tratamiento detallado de la implementación computacional mediante diferentes plataformas de software estadístico reconoce la realidad tecnológica de la práctica veterinaria moderna. La

presentación de código en R, Python, SPSS y SAS no busca convertir al veterinario en programador, sino proporcionar las herramientas prácticas necesarias para implementar los análisis de manera eficiente y reproducible. Esta aproximación multiplataforma reconoce la diversidad de recursos disponibles en diferentes contextos institucionales y geográficos, asegurando que el conocimiento adquirido pueda aplicarse independientemente del software específico disponible. La capacidad de implementar estos análisis de manera autónoma empodera al veterinario para realizar investigación original y análisis de datos sin depender exclusivamente de consultores estadísticos externos. Las aplicaciones en predicción de variables clínicas demuestran el poder transformador de estas técnicas cuando se aplican a problemas específicos de la medicina veterinaria. El modelado farmacocinético permite optimización individualizada de protocolos terapéuticos, considerando las características particulares de cada paciente. La predicción de respuesta a tratamientos oncológicos facilita decisiones informadas sobre opciones terapéuticas y comunicación realista de pronósticos. Los modelos de crecimiento en producción animal permiten optimización económica y mejora de eficiencia productiva. La predicción de fertilidad en reproducción asistida maximiza el retorno de inversión en tecnologías reproductivas costosas. Estas aplicaciones concretas ilustran cómo el dominio de las técnicas de correlación y regresión se traduce directamente en mejor atención clínica, investigación más rigurosa y sistemas productivos más eficientes.

7.8 anexos

7.8.1 Ejercicios Adicionales con Soluciones

Ejercicio Avanzado 1: Diseño de Muestreo Multietápico

Una región ganadera tiene 50 municipios con un total de 2,500 fincas y aproximadamente 100,000 bovinos. Se desea estimar la prevalencia de tuberculosis bovina con precisión del 2% y 95% de confianza. La prevalencia esperada es del 5%.

Solución Detallada:

Paso 1: Calcular tamaño de muestra si fuera aleatorio simple $n = (1.96^2 \times 0.05 \times 0.95) / 0.02^2 = 456$ animales

Paso 2: Diseño multietápico

- Primera etapa: Seleccionar 10 municipios (20% del total)
- Segunda etapa: Seleccionar 5 fincas por municipio = 50 fincas
- Tercera etapa: Muestrear 15 animales por finca = 750 animales

Paso 3: Ajustar por efecto del diseño $DEFF$ estimado $= 1 + (15-1) \times 0.10 = 2.4$ Muestra efectiva $= 750/2.4 = 313$ unidades independientes

Como $313 < 456$, se debe aumentar el número de fincas a 7 por municipio: Total $= 10 \times 7 \times 15 = 1,050$ animales Muestra efectiva $= 1,050/2.4 = 438 \approx 456 \checkmark$

Ejercicio Avanzado 2: Análisis de Variables Complejas

En un estudio de bienestar animal en ponedoras, se registró:

- Densidad de alojamiento (aves/m²): continua
- Sistema de producción (jaula/piso/libre): nominal
- Score de plumaje (0-4): ordinal
- Mortalidad semanal (%): continua
- Presencia de picaje (sí/no): nominal dicotómica

Diseñe un plan de análisis considerando que el objetivo es identificar factores asociados con el picaje.

Solución Detallada:

1. **Variable dependiente:** Presencia de picaje (dicotómica)
2. **Análisis univariado:**
 - Densidad: Comparar medias entre grupos con/sin picaje (prueba t)
 - Sistema: Tabla contingencia, Chi-cuadrado
 - Score plumaje: Mann-Whitney U
 - Mortalidad: Comparar medias o medianas según distribución
3. **Análisis multivariado:**
 - Regresión logística binaria
 - Modelo: $\text{logit}(P(\text{picaje})) = \beta_0 + \beta_1(\text{densidad}) + \beta_2(\text{sistema}) + \beta_3(\text{score}) + \beta_4(\text{mortalidad})$
4. **Verificación de supuestos:**

- Linealidad del logit para variables continuas
- Ausencia de multicolinealidad ($VIF < 5$)
- Observaciones independientes

5. Interpretación:

- Odds ratios con IC 95%
- Ajuste del modelo (Hosmer-Lemeshow)
- Capacidad predictiva (curva ROC)

7.8.2 Anexos técnicos

Anexo 1: Fórmulas de tamaño de muestra

Para estimar una media: $n = (Z_{(1-\alpha/2)} \times \sigma / E)^2$

Para estimar una proporción: $n = Z_{(1-\alpha/2)}^2 \times p(1-p) / E^2$

Para comparar dos proporciones: $n = [Z_{(1-\alpha/2)} \sqrt{2\bar{p}(1-\bar{p})} + Z_{(1-\beta)} \sqrt{(p_1(1-p_1) + p_2(1-p_2))}]^2 / (p_1-p_2)^2$

Para comparar dos medias: $n = 2\sigma^2(Z_{(1-\alpha/2)} + Z_{(1-\beta)})^2 / \delta^2$

Para estudios de concordancia: $n = 8(Z_{(1-\alpha/2)})^2 \times p_0(1-p_0) / d^2$

Anexo 2: Valores críticos comunes

Tabla 47. Valores críticos comunes.

Confianza	α	$Z_{(1-\alpha/2)}$	$t_{((90))}$	$\chi^2_{(0.05,1)}$
90%	0.10	1.645	1.660	2.706
95%	0.05	1.960	1.984	3.841
99%	0.01	2.576	2.626	6.635

Anexo 3: Código R para análisis básicos

```
# Configuración inicial
library(tidyverse)
library(epiR)
library(DescTools)
# Función para análisis descriptivo completo
analisis_descriptivo <- function(datos, variable, grupo = NULL) {
  # Verificar tipo de variable
  if(is.numeric(datos[[variable]])) {
    # Variable cuantitativa
    resumen <- datos %>%
      summarise(
        n = n(),
        media = mean(!sym(variable), na.rm = TRUE),
        mediana = median(!sym(variable), na.rm = TRUE),
        de = sd(!sym(variable), na.rm = TRUE),
        cv = (de/media) * 100,
        min = min(!sym(variable), na.rm = TRUE),
        max = max(!sym(variable), na.rm = TRUE),
        q1 = quantile(!sym(variable), 0.25, na.rm = TRUE),
        q3 = quantile(!sym(variable), 0.75, na.rm = TRUE),
        asimetria = Skew(!sym(variable), na.rm = TRUE),
        curtosis = Kurt(!sym(variable), na.rm = TRUE)
      )
    # Test de normalidad
    normalidad <- shapiro.test(datos[[variable]])
  } else {
    # Variable categórica
    resumen <- datos %>%
      count(!sym(variable)) %>%
      mutate(
        porcentaje = n / sum(n) * 100,
        ic_lower = BinomCI(n, sum(n))[2] * 100,
        ic_upper = BinomCI(n, sum(n))[3] * 100
      )
  }
  return(list(resumen = resumen, normalidad = normalidad))
}
# Ejemplo de uso
# resultado <- analisis_descriptivo(datos_veterinarios, "peso_kg")
```

Figura 71. Código R para análisis básicos.

Anexo 3: Guía rápida de selección de medidas descriptivas

Tabla 48. Guía rápida de selección de medidas descriptivas.

Tipo de Variable	Medida Central	Medida de Dispersión	Gráfico Recomendado
Continua simétrica	Media	Desviación estándar	Histograma
Continua asimétrica	Mediana	RIC	Box plot
Ordinal	Mediana	Rango o RIC	Gráfico de barras
Nominal	Moda	-	Gráfico circular o barras
Conteo	Media o Mediana	Varianza	Histograma o barras
Proporciones	Media	Error estándar	Gráfico de barras con IC
Series temporales	Media móvil	Bandas de confianza	Gráfico de líneas

GLOSARIO

Asimetría: Medida del grado de desviación de la simetría en una distribución de datos.

Box plot: Diagrama de caja que representa visualmente cinco estadísticos resumen y valores atípicos.

Bioestadística: Aplicación de métodos estadísticos a problemas biológicos y médicos, incluyendo medicina veterinaria.

Censura: En análisis de supervivencia, cuando el evento de interés no ha ocurrido al final del período de observación.

Coefficiente de correlación intraclase (ICC): Medida de la similitud entre observaciones dentro del mismo grupo o cluster.

DEFF (Design Effect): Factor que indica cuánto aumenta la varianza del estimador debido al diseño de muestreo complejo comparado con muestreo aleatorio simple.

Error Tipo I (α): Probabilidad de rechazar la hipótesis nula cuando es verdadera (falso positivo).

Error Tipo II (β): Probabilidad de no rechazar la hipótesis nula cuando es falsa (falso negativo).

Inferencia: Proceso de obtener conclusiones sobre una población basándose en información de una muestra.

MCAR (Missing Completely at Random): Datos faltantes cuya ausencia es completamente aleatoria y no relacionada con ninguna variable observada o no observada.

Poder estadístico: Probabilidad de detectar un efecto real cuando existe $(1-\beta)$.

Sesgo: Desviación sistemática de los resultados respecto al valor verdadero.

Coeficiente de variación (CV): Medida de dispersión relativa expresada como porcentaje de la media.

Cuartil: Valores que dividen un conjunto de datos ordenados en cuatro partes iguales.

Curtosis: Medida del grado de concentración de datos alrededor de la media.

Desviación estándar: Raíz cuadrada de la varianza, expresada en las mismas unidades que los datos originales.

Distribución normal: Distribución de probabilidad simétrica en forma de campana.

Error estándar: Desviación estándar de la distribución muestral de un estadístico.

Histograma: Representación gráfica de la distribución de frecuencias de una variable continua.

Intervalo de confianza: Rango de valores que con cierta probabilidad contiene el parámetro poblacional.

Media aritmética: Suma de todos los valores dividida entre el número de observaciones.

Mediana: Valor que divide un conjunto de datos ordenados en dos mitades iguales.

Moda: Valor o valores que aparecen con mayor frecuencia en un conjunto de datos.

Outlier (valor atípico): Observación que se desvía significativamente del patrón general de los datos.

Percentil: Valores que dividen un conjunto de datos ordenados en 100 partes iguales.

Rango: Diferencia entre el valor máximo y mínimo de un conjunto de datos.

Rango intercuartílico (RIC): Diferencia entre el tercer y primer cuartil.

Varianza: Promedio de las desviaciones al cuadrado respecto a la media.

BIBLIOGRAFÍAS

- Al-kassab, M. (2022). The Use of One Sample t-Test in the Real Data. *JOURNAL OF ADVANCES IN MATHEMATICS*, 21, 134–138.
<https://doi.org/10.24297/jam.v21i.9279>
- Alvitez-Temoche, D., Aguila, A., Iarza-Valencia, D., & Calderon, I. (2024). Current trends and spatial-temporal dynamics of veterinary dentistry research: A scientometric study. *Veterinary World*. <https://doi.org/10.14202/vetworld.2024.666-671>
- Ambarcıoğlu, P. (2021a). The use of effect size in veterinary medicine. *Veterinary Journal of Mehmet Akif Ersoy University*, 6(1), 28–33.
<https://doi.org/10.24880/maeuvsd.823493>
- Ambarcıoğlu, P. (2021b). The use of effect size in veterinary medicine. *Veterinary Journal of Mehmet Akif Ersoy University*, 6(1), 28–33.
<https://doi.org/10.24880/maeuvsd.823493>
- Anderson, D. F., Seppäläinen, T., & Valkó, B. (2017, November 1). *Introduction to Probability*. Cambridge Aspire Website; Cambridge University Press.
<https://doi.org/10.1017/9781108235310>
- Apaza Zúñiga, E., Cazorla Chambi, S., Condori Carbajal, C., Arpasi Meléndez, F. R., Tumi Figueroa, I., Yana Viveros, W., Quispe Coaquira, J. E., Apaza Zúñiga, E., Cazorla Chambi, S., Condori Carbajal, C., Arpasi Meléndez, F. R., Tumi Figueroa, I., Yana Viveros, W., & Quispe Coaquira, J. E. (2022). La Correlación de Pearson o de Spearman en caracteres físicos y textiles de la fibra

de alpacas. *Revista de Investigaciones Veterinarias Del Perú*, 33(3). <https://doi.org/10.15381/rivep.v33i3.22908>

Arias, M., Sangrador, o, & Ortega, P. (2020). Comparación de dos medias. Pruebas de la t de Student. *Evidencias en Pediatría*. <https://evidenciasenpediatria.es/articulo.php?lang=es&id=7699>

Astuhuaman, G. G., Cristóbal, O. E. P., & Janampa, M. V. (2018). Aplicación de las pruebas estadísticas de Wilcoxon y Mann-Whitney con SPSS. *Revista de Investigación Multidisciplinaria CTSCAFE*, 2(4), 15–15.

Awe, O. O., Fagbamigbe, A. F., Evans, E., Olorukooba, A. A., Awe, O. D., Adarabioyo, M. I., Ayeni, O. C., & Babalola, B. T. (2022). Building Biostatistics Capacity in Developing Countries. In *Promoting Statistical Practice and Collaboration in Developing Countries*. Chapman and Hall/CRC.

Backhaus, K., Erichson, B., Gensler, S., Weiber, R., & Weiber, T. (2021). Analysis of Variance. In K. Backhaus, B. Erichson, S. Gensler, R. Weiber, & T. Weiber (Eds.), *Multivariate Analysis: An Application-Oriented Introduction* (pp. 147–203). Springer Fachmedien. https://doi.org/10.1007/978-3-658-32589-3_3

Berlanga, V., & Rubio Hurtado, M. J. (2012). Clasificación de pruebas no paramétricas. Como aplicarlas en SPSS. *Articles publicats en revistes (Mètodes d'Investigació i Diagnòstic en Educació)*. <https://diposit.ub.edu/dspace/handle/2445/45283>

Blas Vilca, P. E. (2022). *INFERENCIA ESTADÍSTICA Variables aleatorias discretas y continuas. Distribuciones de probabilidad: Binomial, normal, t, f, χ^2 y de poisson. Distribución uniforme. Distribución normal. Técnicas de muestreo. Distribuciones muestrales: de la media, de una proporción, de la varianza. Estimación de parámetros. Estimación puntual de parámetros. Estimación de parámetros por intervalos. Intervalo de confianza por la media. Intervalo de confianza para la varianza. Intervalo de confianza para la razón de dos varianzas. Intervalo de confianza para la diferencia entre dos medias. Aplicaciones con spss y con otros paquetes estadísticos.* <https://repositorio.une.edu.pe/entities/publication/repositorio.une.edu.pe>

Bologna, E. (2022). *Capítulo 10 Estimación por intervalo | Un Recorrido por los Métodos Cuantitativos en Ciencias Sociales a bordo de R.* <https://estadisticacienciasocialsr.rbind.io/estimaci%C3%B3n-por-intervalo.html>

Burbano, M. J. P., & Pijal, D. C. B. (2021). Importancia de la bioestadística en odontología: Estadística aplicada a la investigación prueba de Kruskal-Wallis. *Universidad y Sociedad*, 13(S3), 275–282.

Calvache, J. A., & Noguera, M. D. (2007). Estimaciones puntuales e intervalos de confianza: Resumen de los resultados de

investigación. *Revista de la Facultad de Ciencias de la Salud Universidad del Cauca*, 9(3), 57–59.

Candia B, R., & Caiozzi A., G. (2005). Intervalos de Confianza. *Revista Médica de Chile*, 133(9), 1111–1115.
<https://doi.org/10.4067/S0034-98872005000900017>

Castañeda, J. A., & Fabián Gil, J. (2004). Una mirada a los intervalos de confianza en investigación. *Revista Colombiana de Psiquiatría*, 33(2), 193–201.

Corcuera Valverde, J. M. (1994). Análisis intrínseco de la estimación puntual. In *Tesis Doctorals—Departament—Estadística. Universitat de Barcelona*.
<https://diposit.ub.edu/dspace/handle/2445/35460>

Costa Silva, A., Braga Costa, E., Dias da Silva, G., & Silva Lacerda. (2024). *Saúde e suas novas perspectivas – Volume 6 – capítulo 9 – Editora Pascal*. Pascal. <https://editorapascal.com.br/saude-e-suas-novas-perspectivas-volume-6-capitulo-9/>

Cuestas, E., Molina Arias, M., & Ochoa Sangrador, C. (2024). Inferencia estadística: Estimación puntual y por intervalos. In *Medicina Basada en la Evidencia: Lo que siempre quiso saber la evidencia aplicada a la práctica clínica sin morir en el intento*. Lúa Ediciones. <https://ri.conicet.gov.ar/handle/11336/268418>

Cummings, C. O., Mitchell, M. A., Perry, S. M., Fleissner, N., Mayer, J., Lennox, A. M., & Johnson-Delaney, C. A. (2022). Bayesian

- Decision Analysis: An Underutilized Tool in Veterinary Medicine. *Animals*, 12(23), 3414. <https://doi.org/10.3390/ani12233414>
- Dugard, P., Todman, J., & Staines, H. (2010). Analysis of variance (ANOVA). In *Approaching Multivariate Analysis, 2nd Edition* (2nd ed.). Routledge.
- Enoyoze, E., Enoyoze, G. E., Enoyoze, E., & Enoyoze, G. E. (2024). Statistical applications in the biomedical sciences: A review. *International Journal of Science and Research Archive*, 12(2), Article 2. <https://doi.org/10.30574/ijsra.2024.12.2.1433>
- Hernández de la Rosa, Y., Hernández Moreno, V. J., Batista Hernández, N. E., & Tejeda Castañeda, E. (2017). ¿Chi cuadrado o Ji cuadrado? *Medicentro Electrónica*, 21(4), 294–295.
- Hothorn, L. A. (2022). *Seven recommendations for alternatives to the common analysis of variance (ANOVA) with application in the life sciences—Using R* (No. arXiv:2205.15866). arXiv. <https://doi.org/10.48550/arXiv.2205.15866>
- Jiang, J. (2022). *Fundamentals of Probability*. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119716822.ch3>
- Kaps, M., & Lamberson, W. R. (2004). Probability. In *Biostatistics for animal science* (pp. 15–25). <https://doi.org/10.1079/9780851998206.0015>
- Karabulut, E. (2022). Conscious and Correct Use of Biostatistical Methods in Medical Researches: From Planning to Reporting the

Results - Part I. *Duzce Medical Journal*, 24(1), 1–6.
<https://doi.org/10.18678/dtfd.1094422>

King, E., Mueller, M., & Dowling-Guyer, S. (2024). Financial fragility and demographic factors predict pet owners' perceptions of access to veterinary care in the United States in: *Journal of the American Veterinary Medical Association* Volume 260 Issue 14 (2022). *JAVMA*.
<https://avmajournals.avma.org/view/journals/javma/260/14/javma.21.11.0486.xml>

Kumamoto, H., Hisano, S., & Takahashi, K. (2021). Constraints on ultra-low-frequency gravitational waves with statistics of pulsar spin-down rates. II. Mann–Whitney U test. *Publications of the Astronomical Society of Japan*, 73(4), 1001–1009.
<https://doi.org/10.1093/pasj/psab058>

Kumar, S., & Yadav, P. K. (2021). Basic Biostatistics for Laboratory Animal Science. In P. Nagarajan, R. Gudde, & R. Srinivasan (Eds.), *Essentials of Laboratory Animal Science: Principles and Practices* (pp. 107–136). Springer. https://doi.org/10.1007/978-981-16-0987-9_6

Lamb, C. R. (2016). Veterinary diagnostic imaging: Probability, accuracy and impact. *The Veterinary Journal*, 215, 55–63.
<https://doi.org/10.1016/j.tvjl.2016.03.017>

Leaverton, P. E., & Zhu, Y. (2025). Biostatistics in Epidemiology. In S. R. Quah (Ed.), *International Encyclopedia of Public Health*

(Third Edition) (pp. 326–338). Academic Press.
<https://doi.org/10.1016/B978-0-323-99967-0.00220-9>

Lima, V., Almeida, M., & Pandolphi, M. (2024). CARACTERÍSTICAS SOCIODEMOGRÁFICAS E CUIDADOS COM ANIMAIS DE ESTIMAÇÃO: UMA ANÁLISE DA AÇÃO SOCIAL “MEDVET NA COMUNIDADE” EM PARINTINS-AM – ISSN 1678-0817 Qualis B2. *Revista ft.* <https://revistaft.com.br/caracteristicas-sociodemograficas-e-cuidados-com-animais-de-estimacao-uma-analise-da-acao-social-medvet-na-comunidade-em-parintins-am/>

Lušničenko, M., Zolova, A., & Gavrilova, A. (2024). Beyond the p-value: How descriptive statistics unlock insights in veterinary research? *German Journal of Veterinary Research*, 4(3), 140–150.
<https://doi.org/10.51585/gjvr.2024.3.0105>

Mata, F. (2022). Linking epidemiology, biostatistics and animal welfare science. *Revista Veterinaria*, 33(2), 220–222.
<https://doi.org/10.30972/vet.3326186>

Mendivelso, F. (2021). Prueba no paramétrica de correlación de Spearman. *Revista Médica Sanitas*, 24(1).
<https://doi.org/10.26852/01234250.578>

Mishra, A., & Khan, M. M. (2024). The science and art of biostatistics: A comprehensive overview. *Asian Journal of Hospital Pharmacy*, 45–51. <https://doi.org/10.38022/ajhp.v4i2.86>

- Molina, M. (2021). Una historia de t y números. La prueba exacta de Fisher. *Revista Electrónica AnestesiaR*, 13(10). <https://doi.org/10.30445/rear.v13i10.1011>
- Molina, M. (2022). Paso a paso. Prueba de la t de Student para muestras independientes. *Revista Electrónica AnestesiaR*, 14(8). <https://doi.org/10.30445/rear.v14i8.1060>
- Molina, M. (2023). Ciencias o letras. *Revista Electrónica AnestesiaR*, 15(5). <https://doi.org/10.30445/rear.v15i5.1136>
- Nave, D., Serrano, L., & Nave, F. (2024). Caracterización clínica de perros ingresados a la Unidad de Bienestar Animal del Ministerio de Agricultura, Ganadería y Alimentación de Guatemala: Clinical characterization of dogs admitted to the Animal Welfare Unit of the Ministry of Agriculture, Livestock and Food of Guatemala. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*, 5(4), 805–820. <https://doi.org/10.56712/latam.v5i4.2293>
- Oñate, V. M., & Morales-Oñate, B. (2023). MTest: Una Prueba bootstrap para Multicolinealidad. *Revista Politécnica*, 51(2), 53–62. <https://doi.org/10.33333/rp.vol51n2.05>
- Ortiz, J. (2020). Una extensión de la prueba de Friedman | Revista Colombiana de Estadística. *Revista Colombiana de Estadística*. <https://revistas.unal.edu.co/index.php/estad/article/view/9797>
- Paredes-Catota, F., Castillo-Hidalgo, E. P., Almeida-Revelo, O. V., & Vilcacundo-Paredes, H. D. (2021). Revista Científica Facultad de

Ciencias Veterinarias. *Revista Científica de la Facultad de Ciencias Veterinarias*, 32, 1–6. <https://doi.org/10.52973/rcfcv-e32146>

Peña, Y. A., & Pérez, A. L. S. (2015). Procedimientos para la estimación por intervalos de confianza en las investigaciones biomédicas. *Revista Cubana de Angiología y Cirugía Vascular*, 16(1), 92–100.

Peñaherrera-Larenas, M. F., Jurado-Izquierdo, Lady, Acosta-Roby, M. G., & Oviedo-Galarza, J. E. (2021). Coeficiente de Correlación de Pearson entre la Canasta y el Salario Básico en el Ecuador. *Polo del Conocimiento*, 6(10). <https://doi.org/10.23857/pc.v6i10.3548>

Piper, S. K., Zocholl, D., Toelch, U., Roehle, R., Stroux, A., Hoessler, J., Zinke, A., & Konietzschke, F. (2023). Statistical review of animal trials—A guideline. *Biometrical Journal*, 65(2), 2200061. <https://doi.org/10.1002/bimj.202200061>

Quicke, D., Butcher, B. A., & Welton, R. K. (2021). Analysis of variance (ANOVA). In *Practical R for biologists: An introduction* (pp. 155–165). <https://doi.org/10.1079/9781789245349.0013A>

Reynolds, M. (2023). Probability and Statistical Principles. In *A Guide to Virology for Engineers and Applied Scientists* (pp. 243–269). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119853169.ch13>

Sarma, K. V. S., Mohan, A., & Vedururu, S. S. (2022). Statistical methods in clinical studies: An overview. *Journal of Clinical and*

Scientific Research, 11(1), 34.
https://doi.org/10.4103/JCSR.JCSR_69_20

Taylor, P., & Meyer, R. E. (2023). Veterinary clinical research or experiments on pets. *Veterinary Anaesthesia and Analgesia*, 50(5), 383–385. <https://doi.org/10.1016/j.vaa.2023.07.006>

Torres, D., & Normando, D. (2021). Biostatistics: Essential concepts for the clinician. *Dental Press Journal of Orthodontics*, 26, e21spe1. <https://doi.org/10.1590/2177-6709.26.1.E21SPE1>

Watty, A. M. D. (2016). *Estadística para veterinarios y zootecnistas*. Editorial Newton Edición y Tecnología Educativa.

Weng, H.-Y., & Messam, L. L. McV. (2025). Reporting and interpreting statistical results in veterinary medicine: Calling for change. *Journal of Veterinary Internal Medicine*, 39(1), e17258. <https://doi.org/10.1111/jvim.17258>

Yupari-Azabache, I., Bardales-Aguirre, L., Rodriguez-Azabache, J., Barros-Sevillano, J. S., & Rodríguez-Díaz, Á. (2021). Factores de riesgo de mortalidad por COVID-19 en pacientes hospitalizados: Un modelo de regresión logística. *Revista de La Facultad de Medicina Humana*, 21(1), 19–27. <https://doi.org/10.25176/rfmh.v21i1.3264>

Zhao, S., Yang, Z., Musa, S. S., Ran, J., Chong, M. K. C., Javanbakht, M., He, D., & Wang, M. H. (2021). Attach importance of the bootstrap t test against Student's t test in clinical epidemiology: A demonstrative comparison using COVID-19 as an example.

Epidemiology & Infection, 149, e107.
<https://doi.org/10.1017/S0950268821001047>

Zweifach, A. (2024). Samples in many cell-based experiments are matched/paired but taking this into account does not always increase power of statistical tests for differences in means. *Molecular Biology of the Cell*, 35(1), br1.
<https://doi.org/10.1091/mbc.E23-05-0159>



Bioestadística aplicada a las ciencias veterinarias, se publicó en el
mes de diciembre de 2025.

ISBN: 978-9907-0-0420-5

Grupo Editorial BLR
Ecuador
Cel: +593 98 320 4362
[https://grupoblir.com/](https://grupoblir.com/publicaciones@grupoblir.com)
publicaciones@grupoblir.com

BIOGRAFÍA DE LOS AUTORES

Víctor Alejandro Bosquez Barcenes:

Licenciado en Informática Educativa, Especialista en Redes y Comunicación de datos, Magister en Informática Empresarial, Doctor en Educación, cursando Doctorado en Ciencias Informáticas, Master en Estadística Aplicada, con una experiencia académica de docente en la Universidad Estatal del Bolívar desde el año 2012, Docente de módulos de Posgrado en diferentes universidades del país. Profesor investigador desde el año 2016 con un total de 24 artículos científicos de impacto mundial y regional, once ponencias de investigación realizadas dentro de distintas instituciones de educación superior, un libro escrito, participación en cuatro proyectos de investigación, par ciego revisor de artículos científicos, director y par académicos de tesis de investigación para la obtención de título de tercer nivel y cuarto nivel.

Joscelito Bolívar Solano Gaibor:

PDr. en Medicina Veterinaria y Zootecnia de la Universidad Agraria de La Habana, Dr. en Ciencias (PhD) Universidad de Oriente Cuba, ha cumplido las siguientes funciones, Director de Relaciones Internacionales UEB, Coordinador de Proyectos Erasmus Mundus con España Universidad de Valladolid (VECCEU) y Portugal Universidad de Porto(BABEL), Jefe del Programa de Avestruces, trabajó en la SENESCYT y Consejo de Educación Superior por cuatro años, Administrador Temporal (equivalente a Rector) de varias universidades, Presidente de la Comisión de Intervención de la Universidad Vargas Torres, Coordinador y Miembro del Comité Científico del Proyecto de indicadores de Sostenibilidad con la Universidad de Almería, España AECID, Coordinador Proyecto Seguridad Alimentaria Universidad de Islas Baleares Mallorca España, Profesor Investigador de la UEB y autor de varios artículos científicos, Coordinador de la UIC de la Carrera de Medicina Veterinaria, actualmente es Director de Posgrado UEB.

BIOESTADÍSTICA APLICADA A LAS CIENCIAS VETERINARIAS

Estimado lector, este libro aborda la necesidad de ofrecer a los estudiantes de Veterinaria una obra integral de estadística y diseño experimental con una aplicación genuina a las ciencias veterinarias, superando la limitación de los textos genéricos. El texto destaca que la alfabetización estadística es crucial para la práctica diaria, ya sea en el diagnóstico clínico, la optimización de la producción animal o la vigilancia epidemiológica.

La estructura avanza desde los fundamentos hasta las técnicas avanzadas. Su característica distintiva es el enfoque práctico mediante el uso del software R, incluyendo código y ejercicios con datos reales, lo cual responde a las demandas del ejercicio profesional contemporáneo. Además, integra consideraciones éticas fundamentales, como el principio de las 3Rs (Reemplazo, Reducción y Refinamiento), en el diseño experimental.

Con una diversidad de ejemplos que cubren desde animales de compañía hasta especies productivas, la obra busca capacitar a los futuros veterinarios para comprender, interpretar y comunicar análisis estadísticos de manera competente, convirtiendo a la estadística en un lenguaje esencial para la medicina basada en evidencia..

Agradecemos a todos los lectores que se acercan a esta obra con ánimo de aprender, aplicar y transformar.



Grupo Editorial BLR
Ecuador
Cel: +593 98 320 4362
[https://grupobl.com/
publicaciones@grupobl.com](https://grupobl.com/publicaciones@grupobl.com)

ISBN: 978-9907-0-0420-5

