

## Aplicación de pruebas estadísticas de distribución de datos y su utilidad en producción animal<sup>1</sup>

Roberto Martínez-López<sup>2</sup>, Liz Mariela Centurión Insaurrealde<sup>3</sup>

### Resumen

**Introducción:** La elección adecuada de herramientas estadísticas para el análisis inferencial de datos, es fundamental en la ciencia. Así, identificar el comportamiento de las observaciones, es indispensable; para seleccionar con la mayor precisión posible, la técnica estadística que conduzca hacia resultados acertados y conclusiones enriquecedoras. **Objetivo:** Se estudió la distribución de datos en bruto y residuales, procedentes de la cría de bovinos y gallinas mediante la verificación de supuestos paramétricos; a su vez, fueron comparados tres métodos estadísticos, por especie zootécnica, discutiendo su plasticidad, ajuste y precisión. **Materiales y métodos:** Se analizaron en bovinos: condición corporal,

peso vivo, longitud de pelo y constantes bioquímicas (calcio, fósforo, magnesio). En aves: peso vivo, ancho de pecho, longitud de muslo, longitud de cresta, presencia de endo y ectoparásitos. Se aplicaron pruebas de normalidad (Shapiro Wilk y Kolmogorov (Lilliefors)) y homogeneidad de varianzas (Levene). Se consideraron en bovinos: ANOVA con Tukey; ANOVA de Welch con Games Howell y Kruskal Wallis con Dunn. En aves: el test Student, con corrección de Welch y Wilcoxon-Mann-Whitney. **Resultados:** Las pruebas de normalidad mantuvieron resultados similares. Se constató diferencia en criterios de decisión entre los análisis inferenciales, para nivel de magnesio y en longitud de muslo. **Conclusiones:** Se recomienda explícitamente, en estudios veterinarios y zootécnicos, con rigor científico, analizar la normalidad y la homogeneidad de varianza,

- 1 Artículo original derivado del proyecto de investigación: Estudio de parámetros adaptativos de los bovinos criollos de los humedales del Ñeembucú y área de influencia de la Universidad Nacional de Asunción, ejecutado entre julio de 2015 y diciembre de 2020. Financiado por el Consejo Nacional de Ciencia y Tecnología, Paraguay, a través del programa PROCIENCIA con recursos del Fondo para la Excelencia de la Educación e Investigación – FEEI
- 2 Doctor y Máster en Zootecnia, Ingeniero Agrónomo. Investigador del Centro Multidisciplinario de Investigaciones de la Universidad Nacional de Asunción. Director del Programa Universitario de Becas para la Investigación, “Andrés Borgognon Montero” (PUBIABM), Paraguay. Correo: robertomartinezlo@vet.una.py. ORCID: 0000-0003-2882-6675
- 3 Licenciada en Ciencias mención Matemática Estadística. Profesor Asistente en la Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Asunción. Correo: lizmariel@facen.una.py. ORCID: 0000-0001-5526-6349

**Autor para Correspondencia:** robertomartinezlo@vet.una.py  
Recibido: 23/05/2023 Aceptado: 26/01/2024

\*Los autores declaran que no tienen conflicto de interés

para identificar y conocer apropiadamente el patrón de comportamiento de los datos provenientes del trabajo, para de esta forma implementar adecuadamente la herramienta estadística inferencial que contribuirá a

discriminar la casualidad y la causalidad en los eventos tratados.

**Palabras clave:** Normalidad, datos, paramétricos, no-paramétricos, zootecnia.

## Application of statistical tests of data distribution and its usefulness in animal production

### Abstract

**Introduction:** The appropriate choice of statistical tools for inferential data analysis is fundamental in science. Thus, identifying the behavior of the observations is essential; to select, with the greatest possible precision, the statistical technique that leads to accurate results and enriching conclusions. **Objective:** The distribution of raw and residual data from cattle and chicken farming was studied by verifying parametric assumptions; In turn, three statistical methods were compared, by zootecnia species, discussing their plasticity, adjustment and precision. **Materials and methods:** The following were analyzed in cattle: body condition, live weight, hair length and biochemical constants (calcium, phosphorus, magnesium). In chickens: live

weight, breast width, thigh length, crest length, presence of endo and ectoparasites. Tests of normality (Shapiro Wilk and Kolmogorov (Lilliefors)) and homogeneity of variances (Levene) were applied. The inferential methods were considered in bovines: ANOVA with Tukey; Welch's ANOVA with the Games Howell test and Kruskal Wallis with Dunn's test. In birds: the student test, with Welch and Wilcoxon-Mann-Whitney correction. **Results:** Normality tests maintained similar results. A difference was found in decision criteria between the inferential analyses, for magnesium level and thigh length. **Conclusions:** It is explicitly recommended, in veterinary and zootecnia studies, with scientific rigor, to analyze the normality and homogeneity of variance, to appropriately identify and know the behavioral pattern of the data coming from the work, in order to properly implement the inferential statistical tool. that will contribute to discriminating chance and causality in the events treated.

**Keywords:** Normality; data; parametric; non-parametric; zootecnia.

## Aplicação de testes estatísticos de distribuição de dados e sua utilidade na produção animal

### Resumo

**Introdução:** A escolha adequada de ferramentas estatísticas para análise

inferencial de dados é essencial na ciência. Assim, identificar o comportamento das observações é fundamental; selecionar, com a maior precisão possível, a técnica estatística que conduza a resultados corretos e a conclusões enriquecedoras. **Objetivo:** Foi estudada a distribuição de dados brutos e residuais da pecuária bovina e de galinhas, verificando pressupostos paramétricos; por

sua vez, foram comparados três métodos estatísticos, por espécie zootécnica, discutindo sua plasticidade, ajuste e precisão. **Materiais e métodos:** Foram analisados condição corporal, peso vivo, comprimento do cabelo e constantes bioquímicas (cálcio, fósforo, magnésio) em bovinos. Nas aves: peso vivo, largura do peito, comprimento da coxa, comprimento da crista, presença de endo e ectoparasitas. Foram aplicados testes de normalidade (Shapiro Wilk e Kolmogorov (Lilliefors)) e homogeneidade de variâncias (Levene). Foram considerados em bovinos: ANOVA com Tukey; ANOVA de Welch com Games Howell e Kruskal Wallis com Dunn. Em aves: teste de Student, com correção de Welch e Wilcoxon-Mann-Whitney. **Resultados:** Os testes de normalidade

mantiveram resultados semelhantes. Foi encontrada diferença nos critérios de decisão entre as análises inferenciais, para nível de magnésio e comprimento da coxa. **Conclusões:** Recomenda-se explicitamente, em estudos veterinários e zootécnicos, com rigor científico, analisar a normalidade e homogeneidade da variância, para identificar e conhecer adequadamente o padrão comportamental dos dados provenientes do trabalho, a fim de implementar adequadamente a ferramenta estatísticas inferenciais que ajudarão a discriminar o acaso e a causalidade nos eventos tratados.

**Palavras-chave:** Normalidade, dados, paramétrico, não paramétrico, zootecnia.

## Introducción

En investigaciones biológicas en general, zootécnicas en específico, es frecuente analizar la influencia de algunos factores sobre el comportamiento de ciertas características que son de interés, principalmente, conocer si la variabilidad observada es producto de un evento típico o deriva de un proceso aleatorio. En ese contexto, es fundamental la consideración de técnicas estadísticas inferenciales adecuadas, las cuales cumplen un papel importante, debido a que permiten conocer, describir y predecir el comportamiento de la población sujeta a estudio, además de presentar la capacidad de discriminar casualidad de causalidad, en el origen de las variables numéricas a analizar.

En esa línea de pensamiento, el aspecto esencial en una investigación con rigor metodológico del ámbito zootécnico, constituyen las herramientas estadísticas escogidas. Así, es importante señalar dos vertientes de esta, la descriptiva y la inferencial. En la primera, interesa generalmente revisar

de manera primaria los datos de una población, permitiendo conocer la estructura y el comportamiento de las observaciones respecto a la característica que está siendo estudiada, siendo en consecuencia sustanciales para detectar la existencia de posibles valores frecuentes o atípicos, por ejemplo. Mientras que, la estadística inferencial permite realizar la estimación de los parámetros de la población, determinando su rango típico de comportamiento, bajo condiciones de estudio y desde las pruebas de hipótesis, discriminar la aleatoriedad o causalidad en los resultados obtenidos, lo que se considera punto central del rigor científico.

Existen distintos métodos estadísticos que pueden emplearse. Sin embargo, la elección de uno o varios de ellos está en función del tema investigado, de los objetivos establecidos, de las hipótesis, del diseño del estudio y de la estructura de los datos (Martínez-López, 2017; Siegel & Castellan, 2015). Algunas técnicas utilizadas para realizar inferencia se basan en modelos paramétricos, los cuales a su vez están sujetos al cumplimiento de

supuestos distribucionales sobre la estimación de los errores, siendo algunos de ellos: la homogeneidad de varianza, la independencia y la normalidad, además de la escala de medición (Fernández-Lozano et al., 2016; Flores-Muñoz et al., 2019; Kim & Park, 2019; Martínez-López, 2017; Siegel & Castellan, 2015).

Existen varios procedimientos que pueden ser aplicados para verificar si la población de la cual derivan las muestras se ajusta a la distribución normal. Según Cabrera et al. (2017), en general no se puede asumir que una prueba presenta mayor potencia que otra, sin embargo, para objetivos específicos dependiendo de lo que se está estudiando, un test puede ser más robusto que otro. En este sentido, algunas pruebas son: Kolmogorov Smirnov, Lilliefors, Anderson Darling, Cramer-von Mises, Jarque Bera, Shapiro Wilk; Shapiro Francia, entre otros (Cabrera et al., 2017; Hadi, 2018; Sánchez-Espigares et al., 2018). Delante de esto, las preguntas: ¿Qué tan útiles son? ¿Cuál es su nivel de importancia y precisión?

El objetivo de este trabajo consistió en estudiar la distribución de datos en bruto y residuales, mediante la verificación de supuestos paramétricos, específicamente la normalidad y homogeneidad de varianzas; a partir de las características concernientes a dos especies zootécnicas totalmente distintas, de gran producción internacional: bovinos y aves; y finalmente, realizar la comparación de los resultados mediante tres técnicas inferenciales en cada especie demostrando la utilidad de las mismas, en estudios con abordaje en el ámbito de la producción animal, entre estos, la obtención de criterios para el manejo de los mismos, selección de individuos superiores y la toma de decisiones en relación a las distintas actividades aunadas en los pilares de la eficiencia.

## Materiales y Métodos

El trabajo se realizó durante el segundo semestre del 2022, en el Departamento Central de Paraguay. Fueron utilizados datos provenientes de los diversos estudios realizados por el grupo de investigación con recursos zoogenéticos del Paraguay, que trabaja en la caracterización genética, productiva, sanitaria y morfométrica, de diferentes especies zootecnicas criadas en zonas de húmedales y semiáridos del territorio. Se consideraron datos reales derivados de sistemas de producción bovina de carne sobre pasturas y cría de pollos rústicos a galpón, para doble propósito (carne y huevos). Se emplearon en el trabajo, datos oriundos de la cría de lotes bovinos pertenecientes a tres razas distintas, afincadas en ganaderías ubicadas en los Húmedales de Ñeembucú y áreas de influencia, en la zona sur de la Región Oriental del Paraguay. Se incluyeron datos de 48 vacas del mismo grupo contemporáneo, separados por razas (16 cada una de las tres razas consideradas), identificados genéricamente como A, B y C, de los que fueron evaluados seis características: condición corporal (CC, escala del 1 al 5); peso vivo de cada bovino (PV, kilogramo), longitud de pelo (LP, centímetro) y tres constantes bioquímicas de interés zootécnico en miligramos por decilitro de sangre: calcio (CA), fósforo (FO) y magnesio (MA). Todas las variables fueron medidas durante un periodo invernal estándar del país (seco y con temperaturas moderadamente bajas).

Fueron abordadas variables morfométricas obtenidas a partir de 26 aves, con edades entre 6 a 18 meses, de las cuales 18 fueron hembras y 8 machos: PV en Kg; ancho de pecho (AP) en cm; longitud de muslo (LM) en cm y longitud de la cresta (LCR) en cm. Igualmente, fue evaluada la presencia de tipos de endoparásitos y

ectoparásitos observadas en esta especie zootécnica. Para el recuento sobre presencia de endoparásito y ectoparásitos en este lote de aves adultas, fueron utilizados 18 animales (de 26), lo cual es representativamente suficiente para conocer concretamente la existencia de estos parásitos, bastante contagiosos en poblaciones cerradas de convivencia permanente, reduciendo el manipuleo animal, ahorrando costos y tiempo; además de proporcionar la información necesaria fundamental para cumplir el objetivo. Cabe resaltar que, estas aves son animales de genética cerrada en situación de riesgo. Todo, desde una población total oscilante entre 80 y 100 aves de genotipo rústico y local.

El supuesto de la normalidad fue verificado mediante las pruebas de Shapiro Wilk y Kolmogorov Smirnov (Lilliefors). Mientras que la homogeneidad de varianza se analizó a través del test de Levene. Para examinar

el efecto de los grupos genéticos de bovinos (A, B y C) en las características evaluadas, se aplicaron tres métodos: (1) ANOVA con el test de Tukey; (2) ANOVA de Welch con la prueba de Games Howell no asumiendo igualdad de varianzas y (3) la prueba de Kruskal Wallis y el test de Dunn. Por otro lado, para el análisis inferencial de variables observadas en aves adultas, se recurrió a: (1) la prueba t de Student, (2) t con corrección de Welch para varianzas desiguales y (3) Wilcoxon-Mann-Whitney. Los datos fueron procesados a través del software R (R Core Team, 2020), bajo una probabilidad de error del 5% (nivel de significancia igual a 0,05).

## Resultados

En las Tablas 1 y 2, se exponen los resultados obtenidos al verificar los supuestos de las herramientas paramétricas, ANOVA (Tabla 1) y, t de Student para las aves (Tabla 2).

**Tabla 1.**

*Verificación de los supuestos de normalidad y de homogeneidad de varianzas de los residuos obtenidos por el ANOVA, en las distintas variables evaluadas en bovinos.*

	Normalidad						Homogeneidad de Varianzas		
	Shapiro Wilk		Kolmogorov Smirnov (Lilliefors)			Levene			
	Estadístico	P-valor	Decisión	Estadístico	P-valor	Decisión	Estadístico	P-valor	Decisión
CC	0,95	0,08	Cumple	0,11	0,08	Cumple	2,24	0,11	Cumple
PV	0,97	0,27	Cumple	0,07	0,68	Cumple	8,12	0,00	No cumple
LP	0,98	0,59	Cumple	0,10	0,20	Cumple	2,49	0,09	Cumple
CA	0,74	0,00	No cumple	0,14	0,01	No cumple	0,43	0,64	Cumple
FO	0,98	0,72	Cumple	0,06	0,82	Cumple	0,83	0,44	Cumple
MA	0,80	0,00	No cumple	0,11	0,08	Cumple	1,72	0,19	Cumple

CC: condición corporal, PV: Peso vivo, LP: longitud de pelo, CA: calcio, FO: fósforo, MA: magnesio. Cumple:  $p > 0,05$  los residuos siguen una distribución normal y se verifica la homogeneidad de varianzas, en caso contrario no cumple.

**Fuente:** Elaborada por los autores.

**Tabla 2.**

Verificación de los supuestos de normalidad de los datos y homogeneidad de varianzas en las distintas variables evaluadas en aves adultas.

SX		Normalidad						Homogeneidad de Varianzas		
		Shapiro Wilk			Kolmogorov Smirnov (Lilliefors)			Levene		
		Estadístico	P-valor	Decisión	Estadístico	P-valor	Decisión	Estadístico	P-valor	Decisión
PV	H	0,95	0,58	Cumple	0,17	0,14	Cumple	15,09	0,00	No cumple
	M	0,95	0,71	Cumple	0,14	0,87	Cumple			
AP	H	0,93	0,22	Cumple	0,15	0,34	Cumple	0,00	0,95	Cumple
	M	0,95	0,73	Cumple	0,17	0,69	Cumple			
LM	H	0,85	0,00	No cumple	0,29	0,00	No cumple	13,71	0,00	No cumple
	M	0,94	0,64	Cumple	0,17	0,65	Cumple			
LCR	H	0,91	0,12	Cumple	0,20	0,03	No cumple	0,49	0,48	Cumple
	M	0,92	0,46	Cumple	0,15	0,81	Cumple			
EN	H	0,76	0,00	No cumple	0,34	0,00	No cumple	0,00	0,95	Cumple
	M	0,68	0,00	No cumple	0,36	0,02	No cumple			
EC	H	0,59	0,00	No cumple	0,43	0,00	No cumple	0,18	0,67	Cumple
	M	0,55	0,00	No cumple	0,47	0,00	No cumple			

SX: Sexo, H: Hembra, M: Macho, PV: Peso vivo, AP: Ancho de pecho, LM: Longitud de muslo, LCR: Longitud de la cresta, EN: Endoparásito, EC: Ectoparásito, Cumple:  $p > 0,05$  los datos siguen una distribución normal y se verifica la homogeneidad de varianzas, en caso contrario no cumple.

**Fuente:** Elaborada por los autores.

Teniendo en cuenta uno de los test de normalidad, Shapiro Wilk (Tabla 1), los resultados evidencian que, solamente en dos características concernientes a constantes bioquímicas, concretamente, calcio y magnesio, los residuos no se distribuyeron de manera normal ( $p \leq 0,05$ ). Mientras que, empleando la prueba de Kolmogorov Smirnov (Lilliefors), no se cumplió la premisa de la normalidad en el nivel calcio encontrado ( $p \leq 0,05$ ). Como puede observarse, en la Tabla 1, las pruebas referidas condujeron a criterios similares, solamente se constató divergencia entre las mismas en el resultado derivado del análisis de magnesio (Shapiro Wilk,  $p \leq 0,05$ ;

Kolmogorov Smirnov,  $p > 0,05$ ). Por otro lado, considerando los resultados observados en aves adultas, las dos pruebas de normalidad arrojaron criterios iguales (Tabla 2), a excepción de la variable LCR evaluada en hembras (Shapiro Wilk,  $p > 0,05$ ; Kolmogorov Smirnov,  $p \leq 0,05$ ). En cuanto a los resultados obtenidos a través del test de Levene, para todas las características evaluadas en bovinos (Tabla 1), excepto el PV, los grupos presentaron varianzas homogéneas ( $p > 0,05$ ). Mientras que, en aves (Tabla 2), mostraron valores de probabilidad menores que 0,05 el PV y la LM.

Analizando los datos obtenidos a partir de tres grupos genéticos bovinos distintos, de manera inferencial, en la Tabla 3 se exponen

los resultados derivados de las pruebas inferenciales.

**Tabla 3.**

*Comparación de grupos genéticos de bovinos (genotipos) respecto a características de interés zootécnico mediante técnicas paramétricas (ANOVA) y no paramétricas (Kruskal Wallis)*

	Prueba	Grupo genético de bovinos (Genotipos)			P-valor
		A	B	C	
		Media $\pm$ SD	Media $\pm$ SD	Media $\pm$ SD	
CC	ANOVA(Tukey)	3,13 $\pm$ 0,53 B	3,25 $\pm$ 0,77 B	2,56 $\pm$ 0,48 A	0,00
	ANOVA de Welch (Games-Howell)	3,13 $\pm$ 0,53B	3,25 $\pm$ 0,77B	2,56 $\pm$ 0,48A	0,00
	Kruskal Wallis (Dunn)	3,13 $\pm$ 0,53B	3,25 $\pm$ 0,77B	2,56 $\pm$ 0,48A	0,00
PE	ANOVA(Tukey)	277,75 $\pm$ 46,70A	309,63 $\pm$ 82,84A	278,13 $\pm$ 32,02A	0,21
	ANOVA de Welch (Games-Howell)	277,75 $\pm$ 46,70A	309,63 $\pm$ 82,84A	278,13 $\pm$ 32,02A	0,36
	Kruskal Wallis (Dunn)	277,75 $\pm$ 46,70A	309,63 $\pm$ 82,84A	278,13 $\pm$ 32,02A	0,38
LP	ANOVA(Tukey)	3,06 $\pm$ 0,33 B	2,93 $\pm$ 0,62 B	1,88 $\pm$ 0,35A	0,00
	ANOVA de Welch (Games-Howell)	3,06 $\pm$ 0,33B	2,93 $\pm$ 0,62B	1,88 $\pm$ 0,35A	0,00
	Kruskal Wallis (Dunn)	3,06 $\pm$ 0,33 B	2,93 $\pm$ 0,62B	1,88 $\pm$ 0,35A	0,00
CA	ANOVA(Tukey)	8,64 $\pm$ 0,47B	7,96 $\pm$ 0,38A	8,83 $\pm$ 0,98B	0,00
	ANOVA de Welch (Games-Howell)	8,64 $\pm$ 0,47B	7,96 $\pm$ 0,38A	8,83 $\pm$ 0,98B	0,00
	Kruskal Wallis (Dunn)	8,64 $\pm$ 0,47B	7,96 $\pm$ 0,38A	8,83 $\pm$ 0,98B	0,00
FO	ANOVA(Tukey)	6,89 $\pm$ 1,51 A	7,05 $\pm$ 1,16 A	7,72 $\pm$ 1,26 A	0,17
	ANOVA de Welch (Games-Howell)	6,89 $\pm$ 1,51A	7,05 $\pm$ 1,16A	7,72 $\pm$ 1,26A	0,18
	Kruskal Wallis (Dunn)	6,89 $\pm$ 1,51 A	7,05 $\pm$ 1,16 A	7,72 $\pm$ 1,26 A	0,23
MA	ANOVA(Tukey)	1,82 $\pm$ 0,17AB	1,72 $\pm$ 0,38A	1,97 $\pm$ 0,12B	0,02
	ANOVA de Welch (Games-Howell)	1,82 $\pm$ 0,17A	1,72 $\pm$ 0,38AB	1,97 $\pm$ 0,12B	0,00
	Kruskal Wallis (Dunn)	1,82 $\pm$ 0,17A	1,72 $\pm$ 0,38A	1,97 $\pm$ 0,12B	0,00

CC: condición corporal, PE: Peso vivo, LP: longitud de pelo, CA: calcio, FO: fósforo, MA: magnesio, SD: desvío estándar. Letras diferentes en una misma fila, indican diferencia significativa entre grupos,  $p \leq 0,05$ .

**Fuente:** Elaborada por los autores.

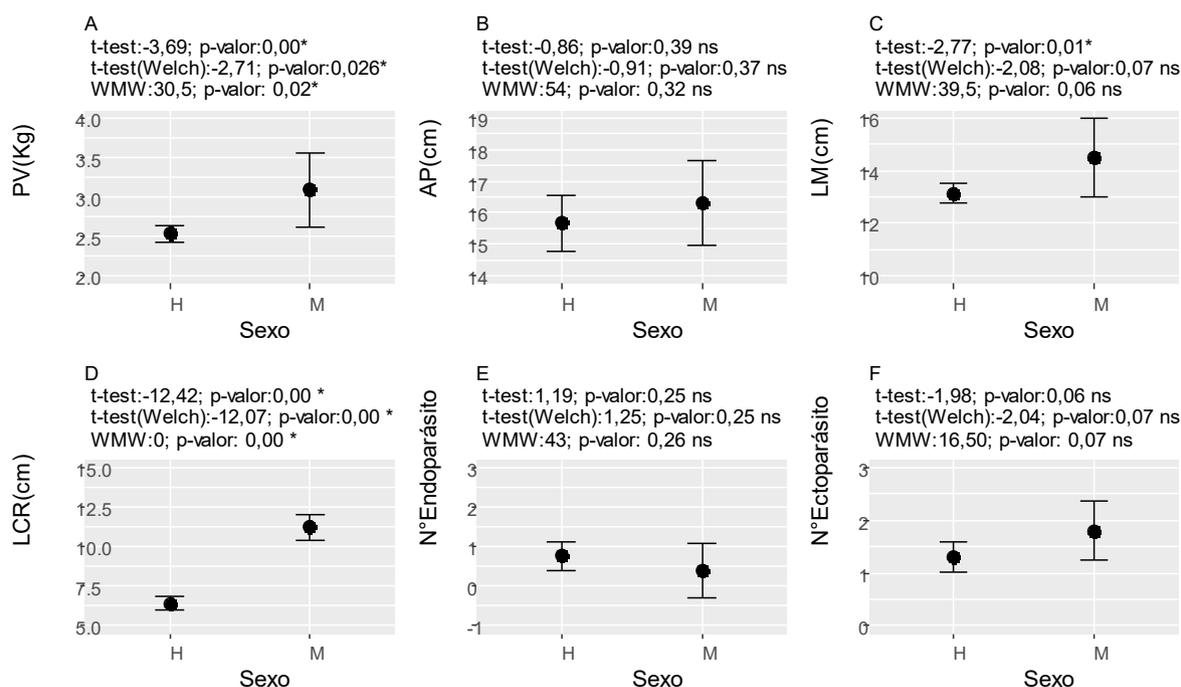
Como puede observarse (Tabla 3), en todas las variables a excepción del nivel de magnesio, las tres pruebas condujeron a mismos criterios de decisión. Cabe resaltar que, tanto el procedimiento paramétrico ANOVA (Tukey), como el ANOVA (Games Howell) no asumiendo igualdad de varianzas y el método Kruskal Wallis (Dunn), arrojaron diferencias en los resultados de la constante bioquímica referida. Para el primer caso, se detectó discrepancia entre el grupo B y C ( $p \leq 0,05$ ), mientras que, en el segundo se observó disimilitud entre A y C

( $p \leq 0,05$ ). Sin embargo, en el tercer y último análisis, C se distinguió significativamente de A y B ( $p \leq 0,05$ ), siendo el genotipo con nivel más alto de magnesio.

En la Figura 1, se muestran los resultados obtenidos mediante la prueba t de Student, t con corrección de Welch y Wilcoxon-Mann-Whitney, al comparar las distintas variables entre aves adultas incluidas en este trabajo, de acuerdo con el sexo.

**Figura 1.**

*Intervalos de confianza al 95% para los valores promedio de las variables medidas en aves adultas. A) PV: Peso vivo. B) AP: Ancho de pecho. C) LM: Longitud de muslo. D) LCR: Longitud de la cresta. E) Cantidad de endoparásito. F) Número de ectoparásito. (\*) diferencia significativa entre grupos  $p \leq 0,05$ , ns: diferencia no significativa. H: hembra; M: macho.*



Fuente: Elaborada por los autores.

Los resultados derivados de las tres pruebas expuestas indicaron mismos criterios para las variables PV (Figura 1A;  $p \leq 0,05$ ), AP (Figura 1B;  $p > 0,05$ ), LCR (Figura 1D;  $p \leq 0,05$ ), Nro. de endoparásito (Figura 1E;  $p > 0,05$ ) y Nro. de ectoparásito (Figura 1F;  $p \leq 0,05$ ), vale decir; diferencia significativa entre machos y hembras en algunas características, mientras en otras, no. Sin embargo, enfatizando la LM, se evidenció discrepancia entre el t-test de Student ( $p > 0,05$ ) y los otros dos, a saber: t-test con corrección de Welch ( $p \leq 0,05$ ), recomendado cuando no se verifica la homocedasticidad (Midway et al., 2020) y, Wilcoxon-Mann-Whitney ( $p \leq 0,05$ ).

## Discusión

Un método útil para explorar el comportamiento de los datos es mediante los gráficos de ilustración. Sin embargo, los mismos no proporcionan evidencia formal y concluyente acerca de la distribución real y precisa de los datos, provenientes de un estudio biológico, en condiciones de campo abierto, donde la variabilidad numérica de los parámetros abordados se estima bastante mayor. Para el efecto, deben considerarse pruebas más específicas (Mishra et al., 2019; Wijekularathna et al., 2022).

Según Cabrera et al. (2017), al momento de escoger el test de normalidad se debe considerar la posibilidad de cometer el error de tipo II, generado al aceptar una hipótesis nula falsa. Igualmente, el nivel de complejidad del procedimiento y la facilidad de su implementación. Al respecto, Hadi (2018) y Wijekularathna et al. (2022) señalan que, las pruebas disponibles y utilizadas generalmente en los paquetes estadísticos son: Shapiro Wilk, Ryan Joiner, Anderson-Darling, Kolmogorov Smirnov (Lilliefors), Jarque Bera, Cramer von Mises. Por los motivos anteriormente señalados, en este trabajo fueron considerados

dos de los test citados, de gran utilización en la comunidad académica y científica en general, en análisis rigurosos de datos (Kim & Park, 2019; Mishra et al., 2019; Vetter, 2017).

Son numerosas las pruebas de normalidad, que pueden ser empleadas en investigaciones. En ese contexto, existen trabajos que abordan el estudio comparativo de potencia y robustez entre los distintos test (Cabrera et al., 2017; Le Boedec, K., 2016; Sánchez-Espigares et al., 2018; Wijekularathna et al., 2022). Algunos autores aducen que, uno de los aspectos a tener en cuenta para la selección adecuada de la prueba, es el tamaño de la muestra (Wijekularathna et al., 2022).

En un estudio sobre el uso de la prueba de normalidad para muestras reducidas, Cabrera et al. (2017) señalaron que, la de Shapiro Wilk presentó mayor potencia que Kolmogorov Smirnov (Lilliefors), es decir, la proporción de veces que el test detectó que en realidad la distribución de los datos no era normal, fue mayor en el primer modelo referido (29 %) en comparación al segundo (17 %), para una muestra de tamaño 15 (probabilidad de error del 5 %). Sin embargo, se registró una leve disminución de potencia y robustez en ambas pruebas, al reducir las observaciones ( $n=10$ ), aunque la asociación relativa se mantuvo. En este sentido, es importante agregar que, en producción animal, trabajar con 10 o menos observaciones, conduciría el análisis hacia resultados bastante frágiles, que muy probablemente desembocaran en conclusiones con sesgo.

De acuerdo con Wijekularathna et al. (2022), evaluando la potencia de 12 pruebas de normalidad, reportaron que el test de Shapiro Wilk funciona moderadamente bien y mejor en el contexto analizado, comparado a Kolmogorov Smirnov (Lilliefors), para un número de observaciones entre 5 y 2000, respectivamente. Al respecto Kim & Park (2019) señalan que, es difícil obtener un

tamaño óptimo de muestra para asegurar la robustez del test de normalidad considerado, no obstante, sugieren considerar una cantidad suficiente de unidades. Por su parte, Le Boedec (2016) indicó que, para muestras reducidas el test puede no ser preciso. En esa línea de análisis, algunos autores refieren que teniendo hasta 40-50 observaciones, se puede emplear el método Shapiro Wilk (Iglesias et al., 2022; Martins-Bessa et al., 2021; Martínez-López, 2017; Mishra et al., 2019). En este trabajo fue considerado el criterio expuesto en esta última alternativa, teniendo en cuenta que el tamaño de muestra, en ambas especies zootécnicas incluidas (bovinos y pollos) fue inferior a 50. Por tanto, la prueba sugerida para analizar la normalidad es Shapiro Wilk.

De acuerdo con lo expuesto por Siegel & Castellan (2015), la potencia y robustez de una prueba se incrementa a medida que el número de observaciones aumenta. Esto suena bastante lógico y puede considerarse de conocimiento común en el ámbito de la ciencia, aunque los costos de experimentación siempre se incrementan con directa proporción, lo cual constituye una fuerte limitante logística y práctica. En este estudio no se evaluó la potencia de las dos pruebas de normalidad empleadas, abordada con métodos de *bootstrap* (Amini-Seresht & Milošević, 2020), debido a que no formaba parte del objetivo, solamente se limitó al empleo y la comparación de los resultados obtenidos analizando algunas variables de interés en la producción animal, para enriquecer el abordaje sobre el asunto puntualmente, retroalimentando la perspectiva de análisis y elección, a la hora de escoger la herramienta primaria de estudio.

Una de las dificultades presentadas al momento de analizar la normalidad es el tamaño reducido de muestra (Cabrera et al., 2017). En el ámbito de la zootecnia no siempre (o casi nunca) es viable considerar una cantidad óptima, puesto que, es imprescindible trabajar con grupos contemporáneos, similares

manejos y genotipos, atendiendo el sexo, finca y sistema agroecológico de cría, entre otros aspectos; que demandan fuerte dificultad logística y/o altos costos operativos, además que en ocasiones son poblaciones de difícil acceso, reducidas o muy grandes, dificultando su cuantificación. Asimismo, dependiendo del objetivo de cada investigación, los animales deben cumplir ciertos criterios de inclusión en la muestra, en este caso en particular para la medición de variables en bovinos, éstos debían encontrarse en estado de reposo, con un ayuno de seis horas, aparentemente sanos, cuyos valores referentes a frecuencias ruminal, cardíaca, respiratoria y temperatura rectal, debían variar dentro del rango natural para la especie, categoría y edad, como también, con peso vivo acorde a los mismos criterios y condición.

En el caso de la variable CC, corresponde a una característica de escala ordinal de 1 a 5, en donde el valor más bajo denota peor condición corporal del animal y el más alto (5), la mejor. Mientras que el recuento de tipo de ectoparásitos y endoparásitos, son de tipo discreto. Según señalamientos en estudios anteriores (Pelea, 2018, 2019), es frecuente un patrón de comportamiento no normal en variables biológicas, debido a la naturaleza de los datos, muchos de ellos medidos en porcentajes, en escala ordinal o de conteo. No obstante, la CC cumplió con el supuesto paramétrico, no así los datos de recuento de endo y ectoparásitos. De acuerdo con Blanca et al. (2017) y Blanca et al. (2023), la prueba F es robusta al incumplimiento de la normalidad. Sin embargo, en un trabajo similar analizando datos en escala de Likert, sin distribución normal, encontraron algunas divergencias entre métodos paramétricos y no paramétricos, en cuanto a los criterios de decisión con respecto a la significancia estadística (Mircioiu & Atkinson, 2017), lo cual coincide con lo evidenciado en el presente trabajo.

En la línea de análisis de la distribución de los datos para las distintas variables medidas en bovinos, se constató que en algunos casos se cumplieron los supuestos de normalidad y homogeneidad de varianzas, permitiendo la consideración de herramientas paramétricas, en este caso, ANOVA. Mientras que, en otros, no se verificó la homogeneidad de varianzas, lo cual es usual trabajando con los datos reales (Blanca et al., 2018). En el contexto presentado y analizado, se puede igualmente aplicar las técnicas citadas anteriormente, pero considerando la corrección de Welch (Delacre et al., 2017; Kim & Cribbie, 2018).

Sobre la comparación de más de dos grupos, como es sabido, el ANOVA no permite identificar entre qué par o pares de estratos existe diferencia significativa. Para ello, se deben emplear pruebas de comparaciones múltiples *a posteriori* o contrastes, al respecto Midway et al. (2020), refieren que existen numerosos casos en donde estos métodos son utilizados inapropiadamente debido a que los usuarios no se encuentran en conocimiento completo sobre las características que informan de su empleo, destacando entre las mismas, la naturaleza de los datos. Según Centurión (2018), analizando investigaciones académicas en el ámbito de la zootecnia, evidenció mayoritariamente el empleo del test de Tukey como procedimiento de comparación por pares. En coincidencia con postulaciones de Midway et al. (2020), quienes aducen el uso frecuente e histórico de la prueba referida. Al respecto, es importante señalar que cuando no es posible asumir igualdad poblacional de varianzas y existen diferencias de los tamaños muestrales entre grupos, es recomendable emplear otras pruebas, tales como: Games Howell, T3 de Dunnett, C de Dunnett, entre otras (Lee & Lee, 2018; Ruffing et al., 2021; Sauder & DeMars, 2019).

En lo que concierne a la prueba *t* de Student, cabe resaltar que, cuando no se cumplen los supuestos paramétricos, la misma puede

conducir a inferencias no válidas (Delacre et al., 2017). En este trabajo, considerando el test referido para la LM, la conclusión podría estar sesgada, asumiendo efecto del sexo de las aves en la LM, cuando en realidad no existe. En ese contexto, teniendo en cuenta los resultados obtenidos al verificar los supuestos, se sugiere la consideración de otras herramientas de análisis, en este trabajo en particular, el test no paramétrico, de acuerdo a las pruebas inferenciales comparadas. Según Fernández-Lozano et al. (2016) y Hopkins et al. (2018), es aconsejable emplear técnicas no paramétricas en el caso planteado; caso contrario, los resultados podrían converger hacia conclusiones equívocas, aumentando la posibilidad de un error falso positivo, pudiendo el investigador concluir, diferencia estadísticamente importante y significativa, entre grupos donde esto realmente no ocurre ni existe (Barrios et al., 2022). Igualmente, pueden ser empleadas otras herramientas en caso de incumplimiento de los supuestos discutidos anteriormente, algunas de ellas son: modelos lineales mixtos, generalizados mixtos (Herrera Villafranca et al., 2020) y los modelos GAMLSS-*generalized additive models for location, scale and shape*- (Hernández Barajas et al., 2021), dependiendo de la naturaleza y objetivo del trabajo abordado.

Haciendo directa alusión a la condición de normalidad y de homocedasticidad, Pelea (2018) refiere que, en investigaciones biológicas se obtienen con frecuencia datos experimentales que no cumplen con las premisas establecidas para poder aplicar un método estadístico paramétrico. De modo general, cuando se incumple alguna o todas las premisas, se suele realizar transformación de escala a los datos, aun así, no siempre se logra el cumplimiento de supuestos. Menciona además otra vía, empleando un método no paramétrico, el cual requiere de un buen diseño del experimento, un tamaño de muestra adecuado (Pataky et al., 2021) y una correcta aleatorización de los tratamientos. No obstante, desde un enfoque

zootécnico se puede aceptar todo lo expuesto aquí, pero se relativiza el punto del “tamaño de la muestra” por razones ya explicadas, que igualmente precisan ser mejor discutidas en otros estudios. De hecho, existen tendencias actuales a trabajar con nuevos modelos de análisis en zootecnia (pasando por alto el tamaño de la muestra), estableciendo caminos mediante la corrección del banco de datos disponibles, utilizando la metodología de redes neuronales artificiales, para pronóstico de curvas de producción de huevos de gallinas (Galeano et al., 2018) o para detectar pérdidas en producción bovina de leche (Borges et al., 2019), sin perder rigor científico. Aquí también deberían ser agregados otros procedimientos complejos, tales como: métodos de remuestreo, aleatorización, permutaciones, enfoques bayesianos, meta-análisis y modelos lineales generalizados (Pelea, 2018).

Siendo un menester, cuando se busca identificar la naturaleza de datos que representa al patrón de comportamiento de un evento biológico en estudio, en el camino de escoger correctamente la inferencia estadística que se utilizará para analizarla, por lo general se intenta direccionar hacia un principio “determinista”. Sin embargo, la naturaleza y su evolución, puede generar patrones bastante dinámicos que generan comportamientos de datos en modo fluctuante, impulsándolo hacia un carácter estocástico. Esto podría deberse a que la evolución en ocasiones se desencadena desde la naturaleza aleatoria de algunos procesos mutagénicos (Molina et al., 2022) y de epistasia (Barton, 2017), fenómenos que también son abordados desde la zootecnia. Por tanto, el carácter determinista, desde la afirmación tajante sobre la convergencia ideal de una herramienta descriptiva que

explique exactamente la fluctuación típica de los datos en estudio, dependiendo qué procesos zootécnicos se analizan, puede originar un sesgo relevante que conduzca a conclusiones inapropiadas.

## Conclusiones

Las dos pruebas de normalidad que fueron evaluadas, mostraron resultados similares en la mayoría de las variables estudiadas, tanto en bovinos como en aves.

Teniendo poblaciones reducidas de aves y de bovinos en estudio, se sugiere primeramente la consideración de la prueba de Shapiro Wilk para conocimiento de la dinámica de comportamiento de los datos a ser analizados. Igualmente, la verificación de la homogeneidad de varianza.

Las herramientas estadísticas inferenciales aplicadas, mostraron distintos resultados en algunos parámetros, llevando en cuenta la probable aceptación y/o rechazo de los criterios analizados, sustentando la verificación de supuestos, entre estos, la elección correcta de las pruebas de normalidad.

Se recomienda explícitamente, en estudios veterinarios y zootécnicos, con rigor científico, analizar la normalidad y la homogeneidad de varianza, para identificar y conocer apropiadamente el patrón de comportamiento de los datos provenientes del trabajo, para de esta forma implementar adecuadamente la herramienta estadística inferencial que contribuirá a discriminar la casualidad y la causalidad en los eventos tratados.

## Referencias

- Amini-Seresht, E., & Milošević, B. (2020). New non-parametric tests for independence. *Journal of Statistical Computation and Simulation*, 90(7), 1301-1314. <https://doi.org/10.1080/00949655.2020.1725007>
- Barrios, Y. D., Guerrero, Z. E., Zambrano, D. F., & Ponce Solórzano, H. X. (2022). Análisis estadístico cuando no se cumplen los supuestos de las pruebas paramétricas, en el contexto de la investigación de la Cultura Física. *Universidad y Sociedad*, 14(S1), 591-600.
- Barton, N. (2017). How does epistasis influence the response to selection? *Heredity*, 118, 96-109. <https://doi.org/10.1038/hdy.2016.109>
- Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: is ANOVA still a valid option? *Psicothema*, 29(4), 552-557. <https://doi.org/10.7334/psicothema2016.383>
- Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2018). Effect of variance ratio on anova robustness: Might 1.5 be the limit? *Behavior Research Methods*, 50(3), 937-962. <https://doi.org/10.3758/s13428-017-0918-2>
- Blanca, M. J., Arnau, J., García-Castro, F. J., Alarcón, R., & Bono, R. (2023). Non-normal data in repeated measures ANOVA: Impact on type I error and power. *Psicothema*, 35(1), 21-29. <https://doi.org/10.7334/psicothema2022.292>
- Borges, P., Mendoza, Z., & Morais, P. (2019). Redes neurales artificiales para estimar las pérdidas en la producción lechera. *Archivos de Zootecnia*, 68(262), 206-212. <https://doi.org/10.21071/az.v68i262.4138>
- Cabrera, G., Zanazzi, J. F., Zanazzi, J. L., & Boaglio, L. (2017). Comparación de potencias en pruebas estadísticas de normalidad, con datos escasos. *Revista de la Facultad de Ciencias Exactas, Físicas y Naturales*, 4(2), 47-52.
- Centurión, L. (2018). *Caracterización de las herramientas estadísticas empleadas en las tesis de grado en el área de la zootecnia* [Tesis de Maestría, Universidad Nacional de Asunción]. [www.conacyt.gov.py](http://www.conacyt.gov.py)
- Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of Student's t-test. *International Review of Social Psychology*, 30(1), 92-101. <https://doi.org/10.5334/irsp.82>
- Fernández-Lozano, C., Gestal, M., Munteanu, C. R., Dorado, J., & Pazos, A. (2016). A methodology for the design of experiments in computational intelligence with multiple regression models. *PeerJ*, 4, 1-22. <https://doi.org/10.7717/peerj.2721>
- Flores-Muñoz, P., Muñoz, L., & Sánchez, T. (2019). Estudio de potencia de pruebas de normalidad usando distribuciones desconocidas con distintos niveles de no normalidad. *Revista Perfiles*, 1(21), 4-11.
- Galeano, L., Galván, I., Aler, R., & Cerón Muñoz, M. F. (2018). Forecasting egg production curve with neural networks. *Archivos de Zootecnia*, 67(257), 81-86.
- Hadi, A. N. (2018). A comprehensive study on power of tests for normality. *Journal of Statistical Theory and Applications*,

- 17(4), 647-660. <https://doi.org/10.2991/jsta.2018.17.4.7>
- Hernández Barajas, F., Cano Urrego, B. & Caicedo Chamorro, E. (2021). Modelos GAMLSS para analizar el grado secado de calcio dihidratado. *Revista EIA*, 18(35), 1-13. <https://doi.org/10.24050/reia.v18i35.1439>
- Herrera Villafranca, M., Galindo Blanco, J., Padilla Corrales, C., Guerra Bustillo, C., Medina Mesa, Y., & Sarduy García, Lucia. (2020). Propuesta del modelo lineal mixto y generalizado mixto para el análisis de un experimento de la microbiología del rumen. *Cuban Journal of Agricultural Science*, 54(2), 149-156. [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S2079-34802020000200149&lng=es&tlng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2079-34802020000200149&lng=es&tlng=es).
- Hopkins, S., Dettori, J. R., & Chapman, J. R. (2018). Parametric and nonparametric tests in spine research: Why do they matter? *Global Spine Journal*, 8(6), 652-654. <https://doi.org/10.1177/219256821878267>
- Iglesias, C., Navas González, F. J., Ciani, E., Camacho Vallejo, M. E., & Delgado Bermejo, J. V. (2022). Bayesian linear regression and natural logarithmic correction for digital image-based extraction of linear and tridimensional zoometrics in dromedary camels. *Mathematics*, 10(19), 1-24. <https://doi.org/10.3390/math10193453>
- Kim, T., & Park, J. (2019). More about the basic assumptions of t-test: normality and sample size. *Korean Journal of Anesthesiology*, 72(4), 331-335. <https://doi.org/10.4097/kja.d.18.00292>
- Kim, Y., & Cribbie, R. (2018). ANOVA and the variance homogeneity assumption: Exploring a better gatekeeper. *British Journal of Mathematical and Statistical Psychology*, 71(1), 1-12. <https://doi.org/10.1111/bmsp.12103>
- Le Boedec, K. (2016). Sensitivity and specificity of normality tests and consequences on reference interval accuracy at small sample size: a computer-simulation study. *Veterinary Clinical Pathology*, 45(4), 648-656. <https://doi.org/10.1111/vcp.12390>
- Lee, S., & Lee, D. K. (2018). What is the proper way to apply the multiple comparison test? *Korean Journal of Anesthesiology*, 71(5), 353-360. <https://doi.org/10.4097/kja.d.18.00242>
- Martins-Bessa, A., Quaresma, M., Leiva, B., Calado, A., & Navas González, F. J. (2021). Bayesian linear regression modelling for sperm quality parameters using age, body weight, testicular morphometry, and combined biometric indices in donkeys. *Animals*, 11(1), 1-23. <https://doi.org/10.3390/ani11010176>
- Martínez-López, R. (2017). *Métodos estadísticos aplicados en Zootecnia*. Etigraf.
- Midway, S., Robertson, M., Flinn, S., & Kaller, M. (2020). Comparing multiple comparisons: practical guidance for choosing the best multiple comparisons test. *PeerJ*, 8, 1-26. <https://doi.org/10.7717/peerj.10387>
- Mircioiu, C., & Atkinson, J. (2017). A Comparison of Parametric and non-parametric methods applied to a Likert scale. *Pharmacy*, 5(2), 26. <https://doi.org/10.3390/pharmacy5020026>

- Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia*, 22(1), 67-72. [https://doi.org/10.4103/aca.ACA\\_157\\_18](https://doi.org/10.4103/aca.ACA_157_18)
- Molina, R. S., Rix, G., Mengiste, A. A., Álvarez, B., Seo, D., Chen, H., Hurtado, J. E., Zhang, Q., García García, J. D., Heins, Z. J., Almhjell, P. J., Arnold, F. H., Khalil, A. S., Hanson, A. D., Dueber, J. E., Schaffer, D. V., Chen, F., Kim, S., Fernández, L. Á., & Liu, C. C. (2022). In vivo hypermutation and continuous evolution. *Nature Reviews Methods Primers*, 2(36), 1-22. <https://doi.org/10.1038/s43586-022-00119-5>
- Pataky T. C., Yagi, M., Ichihashi, N., & Cox P. G. (2021). Landmark-free, parametric hypothesis tests regarding two-dimensional contour shapes using coherent point drift registration and statistical parametric mapping. *PeerJ Computer Science*, 7, 1-25. <https://doi.org/10.7717/peerj-cs.542>
- Pelea, L. P. (2018). ¿Cómo proceder ante el incumplimiento de las premisas de los métodos paramétricos? o ¿cómo trabajar con variables biológicas no normales? *Revista Del Jardín Botánico Nacional*, 39, 1-12.
- Pelea, L. P. (2019). Valores atípicos en los datos, ¿cómo identificarlos y manejarlos? *Revista Del Jardín Botánico Nacional*, 40, 99-107.
- RCore Team. (2020). *A language and environment for statistical computing. R Foundation for Statistical Computing*. Vienna, Austria. <https://www.R-project.org/>
- Ruffing, A. M., Anthony, S. M., Strickland, L. M., Lubkin, I., & Dietz, C. R. (2021). Identification of metal stresses in arabidopsis thaliana using hyperspectral reflectance imaging. *Frontiers in Plant Science*, 12, 1-13. <https://doi.org/10.3389/fpls.2021.624656>
- Sánchez-Espigares, J. A., Grima, P., & Marco-Almagro, L. (2018). Visualizing type II error in normality tests. *American Statistician*, 72(2), 158-162. <https://doi.org/10.1080/00031305.2016.1278035>
- Sauder, D. C., & DeMars, C. E. (2019). An updated recommendation for multiple comparisons. *Advances in Methods and Practices in Psychological Science*, 2(1), 26-44. <https://doi.org/10.1177/2515245918808784>
- Siegel, S., & Castellan, N. (2015). *Estadística no paramétrica: aplicada a las ciencias de la conducta*. Trillas.
- Vetter T. R. (2017). Fundamentals of Research Data and Variables: The Devil Is in the Details. *Anesthesia & Analgesia*, 125(4), 1375-1380. <https://doi.org/10.1213/ANE.0000000000002370>
- Wijekularathna, D. K., Manage, A. B. W., & Scariano, S. M. (2022). Power analysis of several normality tests: A Monte Carlo simulation study. *Communications in Statistics-Simulation and Computation*, 51(3), 757-773. <https://doi.org/10.1080/03610918.2019.1658780>