
Editorial

Isaac Zainea^a
carloszainea@usta.edu.co

Comunicaciones en Estadística presenta el segundo número del volumen 18, correspondiente a diciembre de 2025. En esta entrega se reúnen cinco artículos que reflejan la diversidad de aplicaciones de la estadística y la ciencia de datos en contextos latinoamericanos, abarcando desde el procesamiento de lenguaje natural hasta el diseño muestral en encuestas oficiales.

El número abre con el trabajo de Algarra, León, Camargo y Cruz, quienes abordan la clasificación automática de correos electrónicos no deseados mediante dos arquitecturas de aprendizaje profundo: BERT y LSTM. Los autores evalúan el compromiso entre precisión y costo computacional sobre el Enron Email Corpus, encontrando que ambos enfoques alcanzan una exactitud del 97 %, aunque con diferencias significativas en recursos requeridos. El estudio ofrece orientaciones prácticas para la selección de modelos según las restricciones del entorno de implementación.

A continuación, el artículo de redes neuronales de Elman y modelos GARCH propone una comparación entre técnicas econométricas clásicas y métodos de aprendizaje profundo para el pronóstico del precio de la energía eléctrica en Colombia. Los resultados evidencian que las redes de Elman, tanto en su versión univariada como multivariada, superan al modelo ARMA-EGARCH en términos de error de predicción, lo que resalta el potencial de las redes recurrentes en el modelado de series temporales con alta volatilidad.

El tercer artículo, escrito en inglés por Quintero y Martínez, explora la relación entre las calificaciones internas de una Institución Educativa Distrital y los resultados de las pruebas Saber 11 mediante análisis de correlaciones canónicas y modelos lineales generalizados. El estudio identifica a las asignaturas de Lengua Castellana e Idioma Extranjero como los predictores más fuertes del desempeño en las pruebas estandarizadas, aportando evidencia útil para la toma de decisiones pedagógicas con base en datos institucionales.

En el cuarto trabajo, se presenta un proceso de simulación para evaluar el tamaño de muestra de la Encuesta de Acceso y Uso de los Servicios de Telecomunicaciones en Costa Rica. Mediante 228 escenarios simulados con 5 000 réplicas cada uno, los autores determinan la cantidad óptima de unidades primarias de muestreo por región, logrando una reducción potencial de la muestra sin comprometer la precisión de los indicadores clave. Este artículo constituye un aporte directo a la estadística oficial y al diseño eficiente de encuestas en la región.

El número cierra con el trabajo de Barrera, Ramírez, Lombo y Cruz, quienes aplican un modelo jerárquico bayesiano al análisis del rendimiento futbolístico en la Liga BetPlay 2023-I. El modelo estima simultáneamente los efectos de ataque, defensa y ventaja de localía para los equipos de la liga colombiana, con énfasis en Millonarios F.C. y Atlético Nacional. Los intervalos de credibilidad obtenidos permiten una evaluación probabilística del desempeño deportivo, ilustrando cómo la inferencia bayesiana puede enriquecer la analítica deportiva en contextos poco explorados.

Los cinco artículos de este número comparten un rasgo común: la aplicación de herramientas estadísticas y computacionales a problemas concretos del entorno colombiano y latinoamericano. Desde la detección de spam hasta la optimización de encuestas nacionales, los trabajos aquí reunidos demuestran que la estadística aplicada sigue siendo un puente fundamental entre la teoría y las necesidades de la sociedad.

^aFacultad de Estadística, Universidad Santo Tomás

Invitamos a la comunidad académica a continuar contribuyendo con trabajos que fortalezcan este diálogo entre metodología y aplicación.