



ESTADÍSTICA

# Convertir datos en conocimiento científico

Fue considerada una herramienta pero hoy toma cada vez más valor a partir de las nuevas tecnologías y la acumulación de datos que hacen replantear y ampliar el ámbito de influencia de esta disciplina que se renueva constantemente.

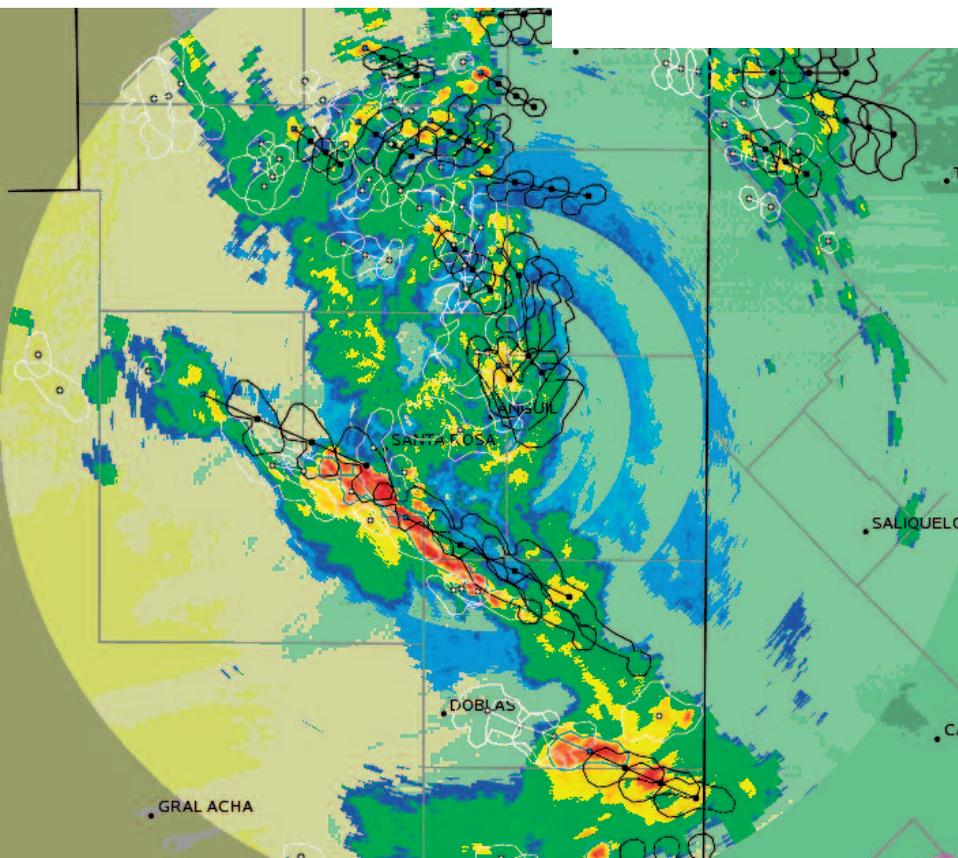
Datos, datos y más datos. Estamos rodeados por ellos, inmersos en ellos, aunque a veces no lo sepamos ni los entendamos. Incluso al abrir esta revista, la mente interpreta los datos que se captan al mirar: números, colores, formas, gráficos y letras. Sin embargo, de acuerdo a cómo son presentados, muestran un tipo de información única. Salvando las distancias (grandes o pequeñas), algo similar proporciona la estadística: ver más allá y extraer información a partir de esos datos que derivarán en conocimiento y viceversa.

“Muchas veces es considerada como una **herramienta** para la investigación pero también es, en sí misma, una **disciplina científica** ya que, para poder coleccionar, analizar e interpretar grandes volúmenes de datos, no sólo hay que usar métodos estadísticos sino también desarrollarlos”, asegura a la Revista RIA la especialista en estadística de la Universidad Nacional de Córdoba (UNC), Mónica Balzarini.

Para Francisco Babinec y Priscila Willems, dos especialistas en estadís-

tica que se desempeñan en el INTA Anguil y Bariloche, respectivamente, “es importante avanzar en su conocimiento y difusión para **favorecer la aplicación de metodologías** alternativas y de avanzada puestas al alcance del investigador y del extensionista”. En este sentido, la estadística interviene desde el momento inicial del trabajo hasta la fase final de análisis, interpretación y elaboración de conclusiones.

Las ciencias agropecuarias la utilizan para generar nuevos conociemien-



**“LA ESTADÍSTICA NOS PROTEGE. LA CONCLUSIÓN ERRADA PROVEE INFORMACIÓN QUE PRONTO SERÁ EXPULSADA DEL CUERPO DE CONOCIMIENTOS CIENTÍFICOS” (MÓNICA BALZARINI).**

tos puesto que requieren de experimentos y estudios. Por ello, “el primer paso de una investigación debe ser una **planificación cuidadosa de la experiencia** en total coherencia con los objetivos planteados, con una diferenciación clara entre ensayos experimentales y estudios observacionales”, aseguran los profesionales. Es que existen procedimientos para tratar e inferir a partir de los datos conseguidos de acuerdo a cómo fueron tomados, por lo que siempre habrá un **protocolo más adecuado que otro**. En los experimentos, el investigador controla ciertas variables o condiciones y puede postular relaciones causa-efecto, mientras que en los estudios observacionales captura los datos de la realidad tal cual son y puede identificar factores o fuentes de variación.

No obstante, si bien en casi la totalidad de las investigaciones agropecuarias se usa la Estadística debido a que los fenómenos en estudio se miden en más de una unidad, también existen ciertos casos que no tienen esa naturaleza ya que son imposibles de repetir en las mismas condiciones pero que son muy importantes para el avance científico como el análisis anatómico

de un tejido vegetal o el estudio cualitativo de una realidad social.

#### **Nuevas tecnologías, nueva Estadística**

A su vez, los nuevos avances tecnológicos en el campo de las telecomunicaciones ocurridos en los últimos años han acelerado ciertos avances científicos. Es que la automatización de procedimientos de captura de información que se presentan en los secuenciadores de ADN, las imágenes satelitales, las estaciones meteorológicas automáticas y los sensores de propiedades del suelo, por ejemplo, generan importantes volúmenes de datos y **nuevos desafíos estadísticos e informáticos** para su almacenamiento, análisis y uso en tiempo real.

Balzarini sostiene que por este motivo “la estadística se **recrea continuamente**: hay que pensar en nuevos métodos, nuevos algoritmos y nuevos protocolos de análisis, porque los que fueron desarrollados para manejar 100 datos, hoy no funcionan bien con miles y cientos de miles”. Pero lo distintivo de esto es que esta abundancia de datos

también llevó a que la Estadística sea considerada una disciplina dinámica y no una herramienta estática.

En ese caso, el conjunto de modelos es tan amplio que se empezó a **agruparla según el tipo de datos** con el que se trabaje. Entre las ramas más comunes en las Ciencias Agropecuarias se encuentran la **Geoestadística**, la **Bioestadística** y el **Data Mining** (este último busca explorar y analizar grandes cantidades de datos para descubrir reglas y patrones significativos a partir de medios automáticos o semiautomáticos).

La primera se especializa en el análisis de conjunto de **datos espaciales o espaciotemporales**. Fue originalmente desarrollada para mejorar las operaciones mineras mediante la observación de la distribución en el espacio geográfico de diversos minerales. En las Ciencias Agropecuarias su utilización está fuertemente ligada con el estudio de datos georreferenciales empleados en la **agricultura de precisión** que, además de información matemática, también utiliza tecnología de avanzada como GPS, SIG, sensores, satélites e imágenes aéreas.

**“UN PASO IMPORTANTE DE UNA INVESTIGACIÓN DEBE SER UNA PLANIFICACIÓN CUIDADOSA DE LA EXPERIENCIA” (BABINEC Y WILLEMS).**

## HISTORIA DE LA ESTADÍSTICA AGROPECUARIA

“Estadística” proviene de “Estado” debido a que era función de éste registrar la población, los nacimientos, las defunciones, los impuestos y las cosechas, entre otros. La finalidad fue tener datos cifrados sobre la población y sus condiciones materiales de existencia.

Si bien los orígenes de la disciplina pueden remontarse al cálculo de probabilidades de Blaise Pascal en el siglo XVII, es con los trabajos de Francis Galton, Karl Pearson y William Gosset (“Student”) a fines del siglo XIX y principios del XX, y por las importantes contribuciones de Ronald Fisher (primera mitad del siglo XX) que se establece la estadística como herramienta fundamental en la investigación científica y tecnológica modernas en el ámbito agropecuario, explican Babinec y Willems, del INTA Anguil y Bariloche, respectivamente.

En 1919 Fisher comenzó a desempeñarse como investigador en la Estación Experimental Agropecuaria de Rothamsted en Inglaterra, donde se dedicó al diseño de experimentos y elaboró sus trabajos sobre el análisis de varianza. El profesional aportó mucho al uso de la estadística en las Ciencias Agropecuarias al introducir y desarrollar ideas originales en el campo de la inferencia estadística y en el de diseño de experimentos.

“Los trabajos posteriores de William Cochran, Gertrude Cox, George Snedecor, Walter Federer, Charles Henderson, David Cox, John Nelder y H. D. Patterson entre muchos otros, le dieron forma a la disciplina tal como la conocemos hoy y en particular a los diseños experimentales de amplia difusión en el ámbito de la investigación agropecuaria”, afirmaron los especialistas.

“El uso correcto de las herramientas de la agricultura de precisión y el manejo de insumos y cultivos según ambientes, permite que los productores sean precisos y eficientes”, indica Andrés Méndez, del proyecto Máquinas y Agrocomponentes Precisos del INTA.

Por otra parte, la segunda rama estudia los problemas planteados dentro de las ciencias de la vida como la **Medicina, la Biología y la Agronomía**. Debido a que los organismos vivos con los que trabaja son muy variados (humanos, animales, plantas, microorganismos, etc.), esta disciplina se encuentra en constante desarrollo. En este sentido, “incluye no sólo herramientas para el análisis estadístico

descriptivo de datos biológicos, sino también el uso de numerosos procedimientos y algoritmos de cálculo y computación para el análisis inferencial, el reconocimiento de patrones en los datos y la construcción de modelos que permitan describir y analizar procesos biológicos de naturaleza aleatoria”, se explica en el libro “¿Qué es la bioestadística?” de la Cátedra de Estadística de la UNC. Esta es una materia que se encuentra en expansión junto con la **Bioinformática** debido a las nuevas biotecnologías y las cantidades de datos biológicos que se generan actualmente.

A su vez, para Babinec y Willems “el mapeo de genes que emplean técni-

cas de biología molecular ha generado inmensas bases de datos cuyo análisis está en continuo desarrollo”. Esto se verá agudizado cuando se profundice con el uso de herramientas como la proteómica, metabolómica, transcriptómica, etc.

Por último, la especialista en **Data Mining** del INTA Anguil, Yanina Bellini, explica a RIA que la aplicación de esa herramienta es “muy reciente” en las Ciencias Agropecuarias en la Argentina. Los desarrollos de modelos se han concentrado en la estimación de superficie, producción o rendimiento de diferentes cultivos y en aplicaciones de la biotecnología. Por ejemplo, en el INTA Corrientes se estudió la generación de



modelos hidrológicos para predecir el caudal de dos ríos de esa provincia con dos meses de anticipación. De esta forma, también se generaron modelos de estimación de granizo en superficie basados en datos proporcionados por radares meteorológicos del INTA, la identificación de sistemas productivos preponderantes en La Pampa y el análisis de micromatrices para caracterizar la coexpresión de genes.

Para Bellini, “la minería de datos nos permite contar con un **nuevo y poderoso set de herramientas** para su análisis e interpretación. Queda claro que la aplicación de estas técnicas, muchas de naturaleza Estadística, sólo se restringe a la disponibilidad de la in-

formación, la imaginación de los especialistas y el trabajo interdisciplinario”.

#### **Mitos y realidades**

Uno pensaría que esta sobreabundancia de datos haría más fácil el desarrollo de la investigación agropecuaria y, de hecho, en cierto aspecto esto es verdad si se considera el uso de las nuevas tecnologías también para su análisis. Según Balzarini, “**hoy se ha complejizado el cálculo** mediante la inclusión de más variables y sus interacciones, pero con los computadores esto ya no es un problema, y los resultados de los estudios se aproximan más a la realidad porque ella es com-

“LA ESTADÍSTICA ES CONSIDERADA UNA HERRAMIENTA PARA LA INVESTIGACIÓN PERO TAMBIÉN ES UNA DISCIPLINA CIENTÍFICA” (MÓNICA BALZARINI).

LOS NUEVOS AVANCES TECNOLÓGICOS GENERAN IMPORTANTES VOLÚMENES DE DATOS Y NUEVOS DESAFÍOS ESTADÍSTICOS E INFORMÁTICOS.

pleja”. Por este motivo, el conocimiento agropecuario de las últimas décadas fue más marcado que en los primeros años del siglo XX (ver recuadro “Historia de la Estadística...”).

Para los especialistas del INTA, esta facilidad que existe en el uso de los nuevos programas estadísticos y la variedad de métodos disponibles puede conducir a **análisis inadecuados o a interpretaciones incorrectas**, por lo que señalan que “muy especialmente, se debe tener un conocimiento mínimo de las características básicas de las metodologías, su ámbito de aplicación y correcta utilización”.

Aún así, modificar los datos a conveniencia personal es un mito muy presente cuando se habla de Estadística, aunque tarde o temprano la tergiversación se pone en evidencia. En este sentido, Balzarini afirma que, debido a que las conclusiones se basan en repeticiones, valores esperados y contemplan los errores aleatorios, la persona que fuerce estas conclusiones **sólo conseguirá hacerlo una vez y la realidad, tarde o temprano, quedará a la vista**. “La Estadística nos protege, puesto que lo que no sea cierto tiene baja probabilidad de repetirse, por lo que la conclusión errada provee información que

pronto será expulsada del cuerpo de conocimientos científicos”, concluye.

#### Más información:

*Mónica Balzarini, balzari@agro.uncor.edu*  
Es Ing. Agr. graduada de la UNC, especialista en Estadística Aplicada y miembro del CONICET. Integró el equipo que desarrolló InfoStat, el primer software de análisis estadístico del país.

*Priscila Willems, pwillems@bariloche.inta.gov.ar*  
Es Estadística graduada de la UNR, y especialista en Biometría y Metodología multivariada.

*Francisco Babinec, babinec.francisco@inta.gov.ar*  
Es ingeniero agrónomo egresado de la UN-LaPlata, Se desempeña en el INTA Anguil y es docente de Estadística en la UN La Pampa.

