

## ANÁLISIS Y COMENTARIOS

# PROBLEMAS COMUNES EN EL ANÁLISIS ESTADÍSTICO DE LA INFORMACIÓN AGRONÓMICA

*Freddy Alemán<sup>1</sup>*

### RESUMEN

**Problemas comunes en el análisis estadístico de la información agronómica** Se discuten algunos de los problemas en el análisis de la información que se genera con la investigación agronómica. Se revisa el papel de las comparaciones múltiples en la investigación agronómica y se dan opciones para enriquecer la información. Se enfatiza sobre la forma correcta de proceder en el análisis de experimentos factoriales y las alternativas existentes para sustituir las comparaciones múltiples. Otro aspecto que se resalta son las inconsistencias existentes en los tratamientos cuantitativos. Se dan sugerencias sobre el análisis e interpretación de dicha información.

### ABSTRACT

**Common problems in the statistical analysis of the agronomic information.** Some of the most common problems in the agricultural information analysis are discussed. The correct form to utilize the factorial experiment analysis and the alternatives to substitute multiple comparisons are emphasized. The role played by the multiple comparisons on the agronomic research is reviewed giving valuable options to improve the obtained information. Inconsistencies obtained from quantitative treatments are also discussed. Recommendations about the correct form to analyze and interpret information are given.



### INTRODUCCIÓN

La investigación agronómica es una necesidad urgente en el ámbito centroamericano. Las autoridades y los investigadores están conscientes de la importancia que tiene la generación de tecnología y del gran significado de la misma en el sector agropecuario. En el pasado se ha debatido ampliamente que la ciencia y la tecnología son motores del desarrollo y que invertir en ellas es una prioridad. La investigación, debería ser una tarea de todos los agrónomos, ya que la misma permite la retroalimentación y la exploración de los últimos adelantos de la Ciencia y la Tecnología. En la actualidad, a pesar de las limitaciones, son muchos los proyectos de investigación agronómica que se ejecutan con la participación de técnicos y profesionales de la agronomía.

A pesar del papel de la investigación agronómica la región centroamericana y de lo consciente que se pueda estar de su importancia, son muchos los problemas que se enfrentan para producir información confiable en el sector agropecuario. Los problemas inician desde la

planificación de la investigación, se continúan durante la ejecución y aún persisten durante la fase de análisis y presentación de los resultados. A esta última fase se refiere el presente trabajo. Camacho y Carbonell (1993) dan una buena descripción de los problemas existentes en las principales etapas del proceso de investigación.

Durante las discusiones en los foros de agronomía, la parte estadística se constituye en un tabú para los investigadores. En más de una ocasión se observa la confusión de los investigadores cuando se formulan ciertas preguntas, sobre la forma correcta de analizar los datos provenientes de sus investigaciones. Bajo dichas circunstancias, el investigador fabrica sus respuestas a preguntas "telegrafiadas". Cuántas veces se ha escuchado por ejemplo, por que utilizó Tukey y no Duncan para la separación de medias? y la respuesta prefabricada es "Tukey es más precisa". Por otro lado, si la pregunta es a la inversa, la respuesta casi siempre es: "se trata de un experimento preliminar ... y durante esta etapa de la investigación no hay que ser exigentes ...". Cuántas veces algún colega ha afirmado lo siguiente: "los resul-

<sup>1</sup> Universidad Nacional Agraria, Managua, Nicaragua, Apartado 453. E-mails: freddy@ibw.com.ni; Freddy.Aleman@evp.slu.se

tados de ese experimento no son confiables por que el coeficiente de variación es demasiado alto” ?

La realidad es que a nivel de los países en el área centroamericana y el Caribe, existen problemas en el análisis de la información generada con las investigaciones, lo cual está llevando a conclusiones erradas. Lo anterior no debe tomarse como una crítica, sino como punto de partida para revisar la metodología y encontrar la solución a dicha problemática. Lauckner (1988) realizó una encuesta que incluyó los números de los años 1980-1986 de la revista *Tropical Agriculture*, y al final concluyó que únicamente el 10% de la muestra, cumplió con los principios del análisis y presentación de los resultados; más enfáticamente expresa, que el uso de la estadística de parte de los autores deja mucho que desear. Al igual que Lauckner, hay muchos otros autores que se han referido a dicha problemática. Lo importante es generar una discusión a lo interno de las instituciones para superar los obstáculos y que permita en el futuro investigación de calidad, con conclusiones válidas para el experimento que se establece. De esa forma, se aprovecha el tiempo y los recursos de forma eficiente.

Por lo tanto, se señalan algunos de los principales problemas que pueden enfrentar los investigadores, en la aplicación de la estadística en la investigación agronómica. También se quiere crear una discusión sobre el tema a lo interno de las instituciones, con otras entidades afines para buscar juntos —investigadores y especialistas en la materia— soluciones a dicha problemática.

### Parte de la problemática

En el presente artículo se trata de resumir algunos de los principales problemas que enfrentan los investigadores cuando se analiza la información proveniente de las investigaciones. En otras oportunidades se abordarán cada uno de los temas por separado proponiendo soluciones y se expondrán otras inconsistencias.

En lo que respecta al análisis de experimentos factoriales (producto de la combinación de dos o más factores). Si el investigador seleccionó más de un factor en el experimento, significa que está interesado en observar el patrón de respuesta de los niveles de un factor, en presencia de los niveles del otro factor; por lo tanto, al momento del análisis debe mantener la estructura de los tratamientos. En estos casos, aún cuando existe interacción entre factores, se comete el error de limitar el análisis al estudio de los efectos principales de cada factor, utilizando comparaciones múltiples para determinar diferencias entre medias. En la práctica, por el contrario, el resultado del efecto principal de un factor puede no

ser de interés, ya que su comportamiento depende de la presencia del otro factor. En otros casos, siempre en presencia de interacción, lo que se hace es incluir en el Análisis de variancia (ANDEVA) los tratamientos cruzados y se combinan en el análisis —de forma inapropiada— los efectos principales de los factores y su interacción. El procedimiento anterior encubre la naturaleza de la respuesta. Como paso final a este procedimiento, se obtienen las medias de tratamientos en una sola columna y se separan las medias utilizando una comparación múltiple. Con este análisis, se pierde la posibilidad de enriquecer las conclusiones, y lo peor, dicho procedimiento puede conducir a conclusiones erradas.

Lo correcto en estos casos es —luego del análisis— determinar si hay presencia de interacción, si existiera, hay que evaluar la respuesta del un factor en cada uno de los niveles del otro factor. La información puede ser desplegada en una tabla de doble entrada, facilitando de esa forma la observación de la respuesta. Las comparaciones múltiples no son apropiadas debido a que en la mayoría de los casos, en experimentos factoriales, existen comparaciones que son sugeridas por la estructura factorial del experimento. Lo correcto es descomponer los grados de libertad y la suma de cuadrado asociada a tratamientos y definir comparaciones de interés, que pueden ser entre los factores por separado o producto de la interacción entre factores.

Como ejemplo, si se desea estudiar el efecto de dos productos herbicidas (A y B) en aplicación pre-emergente y post-temprana, sobre el rendimiento de un cultivo X. Para cumplir los objetivos se establece un experimento factorial (2\*2). Las preguntas involucran el efecto de los herbicidas, el efecto del momento de aplicación y la posible interacción entre productos y momento de aplicación. Las respuestas a dichas preguntas se pueden obtener por medio de la observación en las medias de tratamientos o grupos de tratamientos. Las preguntas que se pueden contestar serían:

1. Existe diferencia entre aplicar el herbicida A y el herbicida B?
2. Existe diferencia entre aplicar los herbicidas como pre-emergentes o como post-tempranos ?
3. La diferencia en el efecto de los herbicidas A y B está en dependencia del momento de aplicación?

Las respuestas a dichas preguntas sería conveniente encontrarlas mediante la descomposición de la variancia y la definición de comparaciones entre tratamientos o entre grupos de tratamientos.

Otro problema común es el análisis de tratamientos cuantitativos (dosis de fertilizantes, dosis de productos químicos, densidades de plantas, distancias de siembra, etc.). Casi siempre, el procedimiento que se sigue es realizar el ANDEVA y luego (muchas veces sin tomar en cuenta el resultado de la prueba de hipótesis) se realizan separaciones de medias, utilizando en la mayoría de los casos la prueba de rangos múltiples de Duncan. En honor a la verdad, no hay nada más inapropiado que una prueba de rangos múltiples para dichos tratamientos.

Cuando se utilizan tratamientos cuantitativos, el objetivo no es obtener la respuesta de un determinado nivel, sino qué se quiere obtener en un universo infinito de niveles. Una vez que se han determinado diferencias significativas entre tratamientos, los niveles evaluados difieren entre sí, pero también todos los posibles niveles que se encuentran entre cada uno de los niveles seleccionados para el experimento. En estos casos lo más adecuado es realizar un análisis de regresión para la variable respuesta o realizar la descomposición de los grados de libertad y la suma de cuadrado de los tratamientos (contrastes lineales de un grado de libertad), con el propósito de determinar el tipo de relación existente entre la variable respuesta y todos los valores que puede tomar el factor cuantitativo. El informe de dicho resultado se realiza por medio de una gráfica que muestre la relación entre la variable dependiente y el factor cuantitativo. Se deben incluir la ecuación y las medidas de precisión.

Un tercer problema común en los informes de investigación es el uso indiscriminado de las separaciones de medias (Duncan, Tukey, SNK, etc.). La popularidad de dichas pruebas se ha originado en que la mayoría de los investigadores las usan y en el hecho de que están disponibles en la mayoría de los paquetes estadísticos. Las comparaciones múltiples son aceptadas únicamente cuando se tienen tratamientos sin estructura o materiales no relacionados (*e.g.*, cultivares, productos químicos, etc.).

La prueba más aceptada en la actualidad es la descomposición de la variancia en comparaciones de un grado de libertad (contrastes ortogonales). Esta prueba permite hacer comparaciones entre grupos de tratamientos, o, entre tratamiento y grupo de tratamientos, y no únicamente entre dos tratamientos (pares de medias). Si en un experimento existen tratamientos con alguna afinidad, es posible juntarlos y realizar compara-

ciones con otros tratamientos (*i.e.*, fertilizantes orgánicos contra fertilizantes químicos). Este tipo de comparaciones son menos frecuentes en los informes científicos, en parte por desconocimiento o por la tendencia de seguir fielmente lo que hacen otros investigadores. Little (1981) mencionó que en cursos de experimentación agrícola que ha impartido en diferentes universidades, todos los participantes tenían conocimiento acerca de las separaciones de medias, pero sólo una mínima parte tenía conocimiento acerca de la descomposición de la variancia para realizar los contrastes ortogonales.

## CONCLUSIONES

No se deben utilizar pruebas de rangos múltiples, para separar las medias de los efectos principales, en experimentos factoriales en los cuales existe interacción entre factores en experimentos que incluyen niveles de factores cuantitativos. Las pruebas de rangos múltiples se deben utilizar únicamente en tratamientos sin estructura o cuando se quiere obtener el tratamiento con mejor respuesta. La descomposición de la variancia que permita comparar tratamientos o grupos de tratamientos debe ser retomada para enriquecer las conclusiones de nuestras investigaciones.

Hasta el momento se han comentado únicamente tres de los problemas más comunes en el análisis de la información agronómica. Se pretende contribuir, en el futuro inmediato, con otras inconsistencias existentes en el análisis de la información agronómica. El propósito debe ser generar resultados confiables en las investigaciones y transmitir la forma correcta de analizar la información de los experimentos de campo, suscita de manera urgente, un foro que permita compartir la experiencia de los investigadores.

## LITERATURA CITADA

- CAMACHO, J. A; CARBONELL, E. A. 1995. Estadística e investigación agraria: Problemas existentes. Investigación agraria 8 (3): 293-309
- LAUCKNER, F. B. 1989. Survey of the use of statistics in agricultural research journals. Tropical Agriculture, 66 (1): 2-7.
- LITTLE, T. M. 1981. Interpretation and presentation of results. Horticultural Sciences, 13 (5): 504-506.