

Aplicación de la estadística, el diseño experimental y software en artículos de la *Revista de Producción Animal*

Roberto Vázquez Montes de Oca, Florentino Uña Izquierdo, Guillermo Guevara Viera, Guillermo Pardo Cardoso y Rafael Avilés Merens

Facultad de Ciencias Agropecuarias, Universidad de Camagüey

roberto.vazquez@reduc.edu.cu

RESUMEN

Dada la importancia cada vez más acentuada de la aplicación de herramientas estadísticas y el diseño de experimentos en las investigaciones, experimentales o no, el Consejo Editorial de la *Revista de Producción Animal*, editada en la Universidad de Camagüey, Cuba, revisó los 638 artículos publicados hasta su XX Aniversario (1985-2005). Se clasificaron fundamentalmente por el uso y tipo de análisis estadístico, diseño experimental y software utilizados. Se compararon los resultados de tres períodos de aproximadamente el mismo tiempo cada uno, para evaluar el uso de esas técnicas en distintos momentos. En cada período se incrementó el uso de software y el nivel de complejidad y actualización de las técnicas analíticas; no así el empleo de diseños experimentales. Aunque es evidente la mejoría, se necesita una mayor exigencia en los requisitos de aceptación y en el arbitraje estadístico. Se propone la superación de los colaboradores en el manejo de paquetes estadísticos y las técnicas analíticas más avanzadas.

Palabras clave: *estadística, diseño experimental, técnicas analíticas, software, artículo científico*

Use of Statistics, Experimental Design, and Software in Articles Published in the *Revista de Producción Animal*

ABSTRACT

Due to the ever increasing importance of applying some statistical tools and the experiment design to research, either experimental or of another kind, the Editorial Board of the *Revista de Producción Animal (Animal Production Journal)*, edited by Camagüey University, reviewed the 638 articles herein published from the first issue in 1985 to that corresponding to the journal's 20th anniversary in 2005. Overall articles were classified according to the use and type of statistical analysis, experimental design, and software. Results from three different periods, each with a similar time span, were compared to assess and evaluate the use of the above mentioned tools at different times. An increase in software application, complexity levels, and analytical techniques updating was evident through each period; however, experimental design use was not satisfactory. Although improvement was apparent, more critical requirements are needed in accepting the submitted manuscripts and in statistical arbitration. Therefore, the Editorial Board recommended to train collaborators in managing both the statistical packages and the most advanced analytical techniques.

Key words: *statistics, experimental design, analytical techniques, software, scientific article*

INTRODUCCIÓN

Cada día los docentes, investigadores y directivos se percatan más de la necesidad de la aplicación de la estadística para demostrar las hipótesis de investigación y argumentar sus decisiones en todas las ramas de la ciencia y la técnica.

En la actualidad, la estadística aplicada no puede verse divorciada, sino integrada con la metodología de la investigación y el diseño experimental, todo lo cual está disponible en los paquetes estadísticos profesionales de mayor circulación mundial en sus últimas versiones.

El objetivo final que sería la aplicación de técnicas analíticas apropiadas en la esfera de la gerencia empresarial, no se cumple ni siquiera en todos los países desarrollados del llamado Primer Mundo. Se considera que el milagro industrial ja-

ponés, iniciado en la segunda mitad del siglo XX, se debió a la creación de una atmósfera favorecedora de la manufactura de productos de alta calidad, éxito atribuido al uso de métodos estadísticos y al desarrollo del pensamiento estadístico entre el personal gerencial (Walpole, 1999). Este mismo autor señala que muchos científicos norteamericanos se quejaron de que en los Estados Unidos de Norteamérica no se ha podido inculcar masivamente, como en los japoneses, el pensamiento estadístico masivamente en los docentes, investigadores y analistas empresariales.

Según criterios de Feinstein citados textualmente por Carrillo (2005) "si uno logra medir lo que está diciendo y lo puede expresar en números, es que sabe lo que dice, pero si no lo puede expresar con números, es que el conocimiento que tiene de ello es escaso e insatisfactorio". El propio autor

resalta la importancia de la estadística en la esfera de la comunicación social, elogia los logros de Japón al respecto y exhorta a su estudio como una necesidad inaplazable.

El Comité Editorial de la *Revista de Producción Animal* (*Rev. prod. anim.*, ISSN 0258-6010), consciente de esta necesidad y con motivo del XX Aniversario de su primera edición, orientó la revisión de los artículos publicados en la misma. Se esperaba confirmar el impacto que en los últimos ocho años ha tenido la influencia de más de cincuenta tesis defendidas en la Maestría de Producción Animal Sostenible, la mayor experiencia y nivel científico del claustro y los colaboradores, así como la mejor organización del arbitraje.

MATERIALES Y MÉTODOS

Se analizaron todos los artículos (638) publicados en la *Revista de Producción Animal*, editada en la Universidad de Camagüey, Cuba, entre los años 1985-2006 los cuales fueron clasificados según:

- Período (1985-1990; 1991-1999 y 2000-2006)
- Especie (bovino, porcino, aves, ovino, otros)
- Perfil (manejo y alimentación, salud animal, genética y reproducción y economía agropecuaria)
- Uso de técnicas analíticas (sí y no)
- Uso de software (sí y no)
- Uso de diseño experimental (sí y no)
- Nivel de complejidad (ninguno, elemental, medio y avanzado)
- Tipo de técnica analítica (ninguna, Anova simple, modelo lineal general (GLM) análisis de frecuencias, comparación de proporciones, estadística descriptiva, T- test, estadística multivariada, regresión lineal simple, estadística no paramétrica y otros tipos de regresión.)
- Tipo de software (ninguno, SPSS, Systat, otros, SAS y Abstat)
- Tipo de diseño experimental (ninguno, total-

mente al azar, bloque el azar, arreglos factoriales y cuadrado latino)

- Origen del primer autor (Habana, Villa Clara, Ciego de Avila, Gramma, profesionales de Camagüey, profesores de la Facultad y extranjeros).

Los criterios para clasificar el nivel de complejidad de las técnicas utilizadas se explican seguidamente:

Elemental: Anova simple, regresión lineal simple, test no paramétrico, T-test, estadística descriptiva, comparación de proporciones y análisis de frecuencias.

Medio: modelos de análisis de varianza y covarianza con efectos fijos, aleatorios y mixtos, con o sin interacciones utilizando el procedimiento GLM disponible en todos los software utilizados.

Avanzado: métodos multivariados como análisis factorial de correspondencia, componentes principales, cluster y Manova y otros tipos de regresiones como modelos ARIMA, no lineales, regresión logística y regresión múltiple en general.

A todas las variables se les aplicó análisis de frecuencia, y en los pares de mayor interés se obtuvo una tabulación curvada o el biplot del análisis factorial de correspondencia simple en base a la significación del test Chi-cuadrado. Todos los análisis estadísticos se realizaron con el SPSS versión 13 (2004).

RESULTADOS Y DISCUSIÓN

Los análisis se efectuaron para el total de artículos (678) y separados según la partición hecha por la variable uso de técnicas analíticas (sí = 511 y no = 127). Se obtuvo amplia información de interés para el Consejo Editorial de la revista, no discutida en su totalidad en el presente trabajo por razones de espacio y objetivos trazados.

Centramos la atención en las variables relacionadas con el uso y tipo de técnica analítica, diseño experimental y software, relacionándolas con los tres períodos analizados en la salida tabular cuando $n = 638$; además se incluyen las variables nivel de complejidad y perfil en la salida gráfica cuando $n = 511$ y $n = 507$, respectivamente, al eliminar del análisis la categoría Economía Agropecuaria por su reducida frecuencia.

En la Tabla 1 se destaca que aproximadamente en el 20 % de

Tabla 1. Frecuencias absolutas y relativas para las variables uso de softwares, técnicas analíticas y diseño experimental

Indicadores	Sí		No	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Uso de técnicas analíticas	511	80,1	127	19,9
Uso de software	201	31,5	437	68,5
Uso de diseños experimentales	192	30,1	446	69,9

Tabla 2. Período de uso del software

Uso de software	Períodos			Total	
	1885/1990	1991/1999	2000/2006		
Sí	Recuento	34	68	99	201
	Frecuencia esperada	72,1	75,3	53,6	201,0
	% software	16,9	33,8	49,3	100,0
	% período	14,8	28,5	58,2	31,5
	% del total	5,3	10,7	15,5	31,5
No	Recuento	195	171	71	437
	Frecuencia esperada	156,9	163,7	116,4	437,0
	% software	44,6	39,1	16,2	100,0
	% período	85,2	71,5	41,8	68,5
	% del total	30,6	28,8	11,1	68,5
Total	Recuento	229	239	170	638
	Frecuencia esperada	229,0	239,0	170,0	638,0
	% software	35,9	37,5	26,6	100,0
	% período	100,0	100,0	100,0	100,0
	% del total	35,9	37,5	26,6	100,0

los artículos publicados, no se utiliza ningún tipo de técnica de análisis estadístico. El alto por ciento de no uso de softwares no es un indicador fiable, pues sucede que en muchas ocasiones no se reporta su uso y en otros casos los cálculos se hacen con calculadoras científicas. Más llamativo es el hecho de que en el 70 % de los artículos no se utilice ningún tipo de diseño experimental, lo cual evidentemente está relacionado con la falta de recursos para el montaje de experimentos, agravada durante el transcurso del Período Especial,¹ que inclinó la balanza hacia el uso de la investigación no experimental.

Aunque el uso de software ha mejorado en el tiempo (Tablas 2 y 3), no ocurre lo mismo con el empleo de las técnicas analíticas y de diseño experimental, cuyo comportamiento en los tres períodos no presentó diferencias significativas según el test de Chi-cuadrado, cuyo valor fue siempre menor de 3,485, $P > 0,175$, cero celdas con una frecuencia esperada menor que 5 y la frecuencia mínima esperada superior a 7. Es importante aclarar que nos estamos refiriendo al uso de técnicas analíticas sin tener en cuenta si son las más adecuadas según el caso, lo cual implicaría un análisis casuístico más detallado y adentrarnos en el campo del metaanálisis. Al respecto Welch (1996) en-

¹ Etapa de agudas dificultades económicas por la desaparición del campo socialista en Europa del Este, agravada por el recrudecimiento del bloqueo aplicado a Cuba por EE.UU.

Tabla 3. Pruebas de Chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	86,768 ^a	2	.000
Razón de verosimilitud	86,182	2	.000
Asociación lineal por lineal	82,144	1	.000
N de casos válidos	638,000		

^a 0 casilla (.0 %) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 53,56.

contró 31,7 % de artículos en *American Journal of Obstetric and Gynecology* con empleo de procedimientos estadísticos inapropiados y otro 28,3 % con uso cuestionado.

Valencia *et al.* (2007) resaltan la necesidad de información sintética y analítica obtenida con métodos, técnicas y diseños correctos. Además, aseguran que el desarrollo de cualquier proyecto de investigación requiere de conocimiento elemental sobre estadística para que reporte los beneficios esperados.

En la Tabla 4 se destaca el uso de Anova simple que se corresponde con el diseño completamente al azar (Tabla 5). En general predominan las técnicas más elementales y el poco uso de las más avanzadas.

Entre el análisis de frecuencias y estadística descriptiva sumaron 76 artículos (12 %) que aportan muy poco conocimiento científico. En este sentido, la guía de estudio para los colaboradores de la revista *Annals of Tourism Research* (2004) expresa que los originales meramente descriptivos que no contribuyen al desarrollo del conocimiento

no se consideran aptos para su publicación; por ejemplo el volumen 6 número 2 de esta revista, publica nueve artículos y en todos se aplican técnicas multivariadas de avanzada, incluido el análisis estructural en tres de ellos.

Con respecto al tipo de software utilizado, la Tabla 6 muestra la mayor frecuencia del SPSS (Statistical Package for the Social Science) y el Systat (Statistical System) pero sobre todo resalta el poco uso del SAS (Statistical Analysis System) que es el paquete estadístico modular insignia en la investigación científica, y encima su empleo ha sido con modelos muy sencillos en el PROC GLM, únicamente clasificado en el nivel medio de complejidad (Tabla 7).

En las Figuras 1 y 2 se aprecia cómo el nivel de complejidad avanzado referido al uso de técnica y diseños más actualizados, no se asocia bien con ninguna de las dos categorías o niveles de las variables perfil y periodo; sin embargo, es evidente la relación entre el nivel de complejidad elemental con el perfil salud animal y nutrición y alimentación; así como el nivel de complejidad medio con el perfil de genética y reproducción. También se define con claridad la tendencia lineal de mejora en el tiempo, al observar la distribución de las variables nivel de complejidad y período.

Abundando en el tipo de software utilizado y su relación con el nivel de complejidad, a manera de comparación tomemos como referencia, no una revista como la *Animal Production* o la *Animal of Dairy Science*, sino una de menor impacto, con perfiles afines a los nuestros (nutrición y alimentación, reproducción, calidad de los productos, patología animal y economía y gestión) la Revista ITEA (Información Técnica Económica Agraria) de la Asociación Interprofesional para el Desarrollo Agrario, de Zaragoza, España.

Se revisaron 132 artículos de la VII Jornada sobre Producción Animal (1997) y 144 de la oncena edición (2005) para hacerlos corresponder aproximadamente con nuestros dos últimos períodos de estudio, y utilizando criterios similares en

Tabla 4. Frecuencia absoluta y relativa de los tipos de técnicas analíticas utilizadas

Tipo de técnica	Frecuencia	Porcentaje
Anova (simple)	259	40,6
Ninguna	127	19,9
Modelo lineal general (glm)	62	9,7
Análisis de frecuencias	55	8,6
Comparación de proporciones	41	6,4
Estadística descriptiva	21	3,3
T-test	20	3,1
Estadística multivariada	19	3,0
Regresión lineal simple	15	2,4
Estadística no paramétrica	11	1,7
Otros tipos de regresión	8	1,3
Total	638	100,0

Tabla 5. Frecuencia absoluta y relativa de los tipos de diseños experimentales utilizados

Diseño experimental	Frecuencia	Porcentaje
Ninguno	446	69,9
Totalmente al azar	140	21,9
Bloque al azar	31	4,9
Arreglos factoriales	14	2,2
Cuadrado latino	7	1,1
Total	638	100,0

Tabla 6. Frecuencia absoluta y relativa de los tipos de software utilizados

Software utilizados	Frecuencia	Porcentaje
Ninguno	437	68,5
SPSS	51	8,0
Systat	47	7,4
Otros	46	7,2
SAS	31	4,9
Abstats	26	4,1
Total	638	100,0

Tabla 7. Frecuencia absoluta y relativa para la variable nivel de complejidad

Nivel de complejidad	Frecuencia	Porcentaje
Ninguno	127	19,9
Elemental	422	66,1
Medio	61	9,6
Avanzado	28	4,4
Total	638	100,0

cuanto al nivel de complejidad. Los resultados se observan en la Tabla 8; la diferencia con el 100 % se debe a los artículos que no usaron tipo alguno de software y de ningún nivel de complejidad, que representan el 18,9 y el 4,2 % en los años 1997 y 2005 por su orden.

En el año 1997, los tres niveles de complejidad de esta revista son similares y usan el SAS aproximadamente en el 2 % de los artículos, mientras que ya en el año 2005 se manifiesta la

tendencia a mejorar y el uso del SAS, no solo se incrementa al 51,4 %, sino que se incorporan otros procedimientos del módulo SAS/STAT como PROC NLIN, PROC REG y CATMOD al procedimiento GLM más avanzado en 1997.

También se destaca, en ambos años de esta revista, el uso casi generalizado del diseño experimental, algo no característico en las revistas de corriente principal en la actualidad, donde la tendencia es hacia el estudio de fenómenos complejos y con investigación no experimental.

En general en *Rev. prod. anim.* se nota la ausencia de técnicas y diseños aplicables a sus perfiles, como los módulos de regresión Logia, Probit, Tubit, de Cox, Curra ROC, análisis de supervivencia paramétricos y no paramétricos, series temporales, análisis factorial de correspondencia, diseños intra-sujetos con medidas repetidas, Structural Equation Modeling (SEM) y redes neuronales.

CONCLUSIONES

Evidentemente con el tiempo ha mejorado el uso de software y el nivel de complejidad de las técnicas analíticas empleadas, que aunque todavía insuficientes, han ayudado a elevar la imagen y calidad de la *Revista de Producción Animal*.

Se requiere una mayor exigencia en los requisitos de aceptación y arbitraje estadístico para reducir al mínimo, no solo aquellos artículos que no utilizan ningún diseño, técnica analítica o software, sino incluso aquellos que emplean métodos totalmente descriptivos.

RECOMENDACIONES

Se propone un plan de superación, dirigido a los autores, sobre manejo de paquetes estadísticos, incluido el lenguaje de programación y procedi-

Tabla 8. Comportamiento del nivel de complejidad y tipo de software utilizado en la Revista ITEA

Año	N	Nivel de complejidad (%)			Tipo de software (%)	
		Elemental	Medio	Avanzado	SAS	Otros
1997	132	24,2	31,1	25,8	19,7	61,4
2005	144	7,6	53,5	34,7	51,4	44,4

Fig. 1. Biplot (nivel de complejidad x perfil)

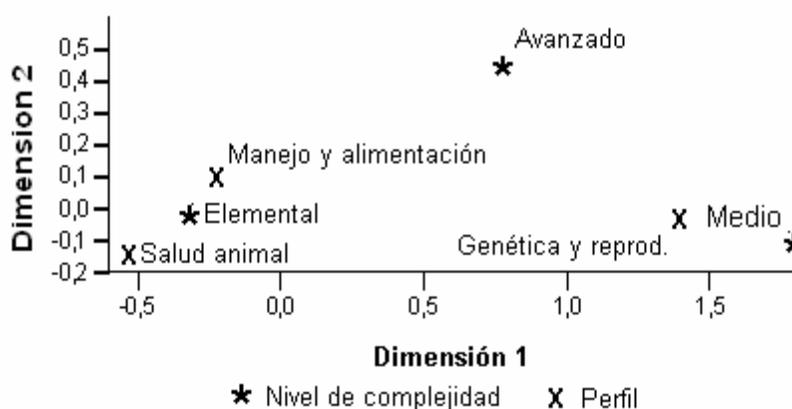
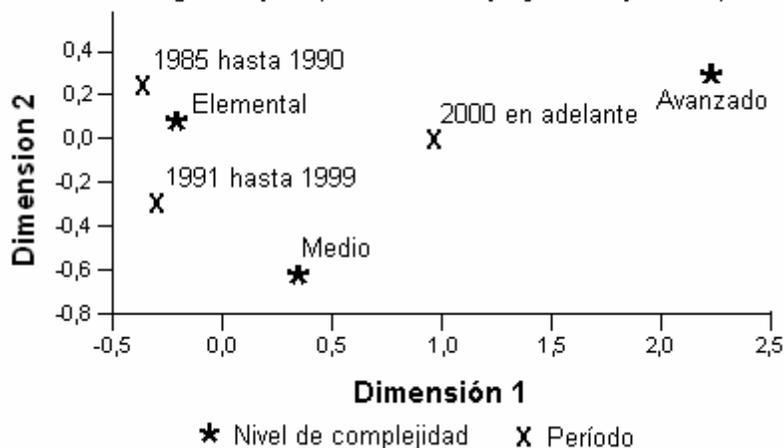


Fig. 2. Biplot (nivel de complejidad x período)



mientos de SAS, así como de técnicas analíticas de avanzada que aparecen en sus últimas versiones para Windows.

REFERENCIAS

CARRILLO, S. R.: "Metodología de la investigación y estadística para comunicadores", *Espacio*, Revista de la ACCS, 10 (abril): 25-28, 2005. (ISSN: 10250239)
 "Guía de estudio para los colaboradores", *Annals of Tourism Research* (en español), Zaragoza, España, 6 (2), 2004. (ISSN: 1575-443-x)
Información Técnica Económica Agraria, "VII Jornada sobre Producción Animal", *ITEA*, vol. extra, 2 (18),

1997. (ISSN:1130-6009. Se consultó el número completo.)
Información Técnica Económica Agraria, XI Jornadas sobre Producción Animal” ITEA, vol. extra, II (26), 2005. (ISSN:1130-6009. Se consultó el número completo.)
SPSS: Statistical Package for the Social Science, 13.0 for Windows, Copyright © SPSS Inc. 1994-2000.
VALENCIA, L. L.; A. S. MUNTANE, A. MASUET, A. V. CALDERÓN Y E. M. L. MORO: “Introducción a la neuroinvestigación”, *Archivos de Neurociencias*, México, 12 (1): 62-67, 2007. (ISSN 0187-4705)
WALPOLE, R.E.: *Probabilidad y estadística para ingenieros*, 6ta. ed., Paerson Education, Prentice-Hall Hispanoamérica, México, 1999.
WELCH, I.; E. GERALD Y G. GABBE: “Review of Statistics Usage in the *American Journal of Obstetric and Gynecology*”, *American Journal of Obstetric and Gynecology*, 175 (5). Nov., 1996.

Recibido: 15/7/2007
Aceptado: 20/8/2007