Monografía | Monograph

El Modelo Lineal General como Tema I para la asignatura Diseño Experimental (Asignatura Optativa I) en la carrera de Ingeniería Agrícola

The General Lineal Pattern as Topic I for the subject Experimental Design (Optional Subject I) in the career of Agricultural Engineering

MSc. José Antonio Pino Roque1, Dra.C. Mayra Arteaga Barrueta2, Dra.C. Lucía Fernández Chuairey1, MSc. Nelson Ulises Lim Chang1, MSc. Vilma Toledo Dieppa1, Lic. Arturo Díaz Gil1

|  |
| --- |
| 1 Departamento Matemática Física, Facultad de Ciencias Técnicas  2 Departamento Química, Facultad de Agronomía  Universidad Agraria de La Habana “Fructuoso Rodríguez Pérez”. Autopista Nacional, carretera Tapaste, km 23 ½, San José de Las Lajas, Mayabeque. |

Autores para correspondencia: pino@unah.edu.cu; mayra@unah.edu.cu; lucia@unah.edu.cu; [limc@unah.edu.cu](mailto:limc@unah.edu.cu); vilma@unah.edu.cu; [arturo@unah.edu.cu](mailto:arturo@unah.edu.cu)

Resumen

La asignatura optativa Diseño Experimental consta de dos temas de estudio. Este material proporciona apuntes de contenidos con el objetivo de facilitar a los estudiantes poder recibir la docencia sobre el Tema I de Modelo Lineal General. Los 5 ejemplos que se brindan están relacionados con el perfil de la carrera de Ingeniería Agrícola y vienen acompañados de orientaciones para su solución y el manejo del paquete estadístico Statgraphics.

**Palabras Claves:** Estadística, diseño experimental, modelo lineal general, regresión, regresión múltiple.

Abstract

The subject optional Experimental Design consists of two study topics. This material provides notes of contents with the objective of facilitating the students to be able to receive the contents on the Topic I of General Lineal Model. The 5 examples that you toasts are related with the profile of the career of Agricultural Engineering and they come accompanied by orientations for their solution and the handling of the statistical package Statgraphics.

**Key words:** Statistic, experimental design, model lineal general, regression, multiple regression.

Recibido: 21 de octubre de 2020

Aprobado: 30 de noviembre de 2020

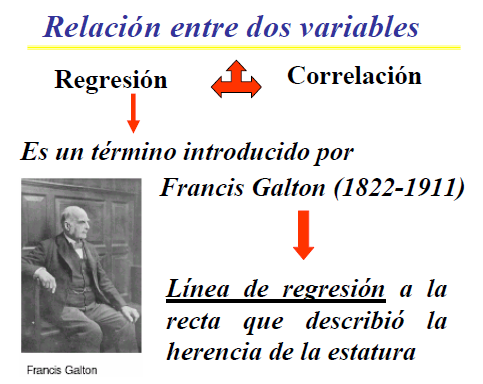
|  |  |
| --- | --- |
| Introducción  La Estadística es la rama de las matemáticas que se ocupa de reunir, organizar y analizar datos numéricos y que ayuda a resolver problemas como el diseño de experimentos y la toma de decisiones.  Di Rienzo *et al*. (2005) señalan que la enseñanza de la Estadística en las ciencias agropecuarias no es un tributo a la modernidad sino una larga tradición que se origina en los trabajos de Fisher que, a comienzos del siglo XX, sentaron las bases de la estadística aplicada a la experimentación agrícola.  Kuehl (2001) plantea en sus análisis históricos que nadie ha tenido tanto impacto en los principios estadísticos del diseño de experimentos en su tiempo como Ronald A. Fisher.  El **Diseño Experimental** consiste enplanificar los pasos a seguir para desarrollar un experimento para la conducción y análisis del experimento de manera tal que se reúna toda la información necesaria para cumplimentar, objetivos de una investigación y que a la vez garantice la precisión y exactitud deseada en los resultados.  Según Guerra et al. (2006) son varios los autores que han tratado de definir o dar una idea clara del significado de lo que se debe entender por diseñar un experimento.  Para Montgomery (1991) un experimento es una prueba o ensayo (literalmente), pero un experimento diseñado es una prueba o serie de pruebas en las cuales se introducen cambios deliberados en las variables de entrada de un proceso o sistema, de manera que sea posible observare identificar las causas de los cambios en las respuestas de salida.  Con esta asignatura, de carácter optativa, se pretende preparar a los estudiantes para que puedan aplicar el modelo lineal general a situaciones de regresión y de análisis de Varianza relacionadas con su perfil profesional, y utilizar los diseños experimentales en situaciones relacionadas con su perfil profesional y un paquete estadístico de cómputo, interpretando las salidas correspondientes.  En este trabajo se abarcan aspectos sobre el Modelo Lineal General ya que en la mayor parte de los problemas de investigación donde se aplica el análisis de regresión se necesita más de una variable independiente en el modelo de regresión.  Es por ello que Walpole *et al*. (2008) consideran que la complejidad de la mayor parte de los mecanismos científicos es tal que para ser capaces de predecir una respuesta importante se necesita un modelo de regresión múltiple.  Es por ello que cuando este modelo es lineal en los coeficientes se denomina modelo de regresión lineal múltiple, aspecto se tratan en este material y que se detalla a través de 8 ejemplos específicos.  Desde el punto de vista educativo se pretende con esta asignatura reafirmar la convicción sobre la materialidad del objeto de estudio de los Diseños Experimentales a través de los modelos estadísticos correspondientes como un reflejo ideal y simplificado de los fenómenos aleatorios que representan una situación real dada. Además, formar un sistema de conocimientos y habilidades de carácter profesional y científico-técnico al aplicar los modelos estadísticos de manera independiente y creadora a la solución de problemas concretos de la especialidad mediante los diseños experimentales correspondientes.  La informática ha venido a aportar la posibilidad de suavizar extraordinariamente la aplicación de los métodos probabilísticas - estadísticos, de modo que ya no se requiere, necesariamente, dominar el aparato de cálculo y procesamiento numérico, sino solamente interpretar los resultados del método estadístico aplicado. Se podría decir que basta, en lo fundamental, con dominar el aspecto cualitativo, el cuantitativo es asunto de la computadora con el software (paquete estadístico) adecuado y, por supuesto, de los estadísticos matemáticos (Miller *et al*., 2007).  El colectivo docente de Estadística Matemática de la UNAH recomienda el uso del paquete estadístico **Statgraphics** **plus 5.1** ya que proporciona facilidades de uso y reúne las condiciones de trabajo según los contenidos que reciben los estudiantes en la Facultad de Ciencias Técnicas. Existen otras versiones que también pueden usarse. La interpretación de cada resultado en el contexto de la problemática y de la base de datos ofrecida es vital para los estudiantes.  Desarrollo  Sobre el Modelo del profesional: aspectos fundamentales  Al consultar la Enciclopedia Cubana de Ecured (Figura #1) sobre el Modelo del Profesional en la carrera de Ingeniería Agrícola se plantea que el egresado de estos estudios es un profesional que está orientado a planificar, gestionar, diseñar, evaluar y supervisar los proyectos de ingeniería, dirigidos tanto a promover el desarrollo social y productivo de los sectores agrícola, forestal, pecuarios, agroindustrial y energético, como a resolver problemas de conservación y aprovechamiento de los recursos naturales y del ambiente. | Introduction  Statistics is the branch of mathematics that deals with gathering, organizing and analyzing numerical data and that helps solve problems such as the design of experiments and decision making.  Di Rienzo et al. (2005) point out that the teaching of Statistics in the agricultural sciences is not a tribute to modernity but a long tradition that originates in the works of Fisher that, at the beginning of the 20th century, laid the foundations of statistics applied to agriculture experimentation.  Kuehl (2001) argues in his historical analyzes that no one has had as much impact on the statistical principles of design of experiments in his time as Ronald A. Fisher.  The **Experimental Design** consists of planning the steps to follow to develop an experiment for the conduct and analysis of the experiment in such a way that all the necessary information is gathered to fulfill the objectives of an investigation and that at the same time guarantees the precision and accuracy desired in the results.  According to Guerra et al. (2006) there are several authors who have tried to define or give a clear idea of ​​the meaning of what should be understood by designing an experiment.  For Montgomery (1991) an experiment is a test or trial (literally), but a designed experiment is a test or series of tests in which deliberate changes are introduced in the input variables of a process or system, so that it is possible I will observe to identify the causes of the changes in the output responses.  This optional subject is intended to prepare students so that they can apply the general linear model to situations of regression and analysis of Variance related to their professional profile, and to use experimental designs in situations related to their professional profile and a statistical computing package, interpreting the corresponding outputs.  This work covers aspects of the General Linear Model since in most of the research problems where regression analysis is applied, more than one independent variable is needed in the regression model.  That is why Walpole et al. (2008) consider that the complexity of most scientific mechanisms is such that a multiple regression model is needed to be able to predict an important response.  That is why when this model is linear in the coefficients it is called a multiple linear regression model, an aspect that is dealt with in this material and which is detailed through 8 specific examples.  From the educational point of view, this subject is intended to reaffirm the conviction about the materiality of the object of study of Experimental Designs through the corresponding statistical models as an ideal and simplified reflection of the random phenomena that represent a given real situation. In addition, to form a system of knowledge and skills of a professional and scientific-technical nature by applying statistical models independently and creatively to the solution of specific problems of the specialty through the corresponding experimental designs.  Computer science has come to provide the possibility of extraordinarily smoothing the application of probabilistic-statistical methods, so that it is no longer necessary, necessarily, to master the apparatus of calculation and numerical processing, but only to interpret the results of the applied statistical method. It could be said that it is enough, fundamentally, to master the qualitative aspect, the quantitative is a matter of the computer with the appropriate software (statistical package) and, of course, of the mathematical statistics (Miller et al., 2007).  The teaching group of Mathematical Statistics of the UNAH recommends the use of the statistical package Statgraphics plus 5.1 as it provides ease of use and meets the working conditions according to the content received by students in the Faculty of Technical Sciences. There are other versions that can also be used. The interpretation of each result in the context of the problem and the database offered is vital for the students.  Development  About the Professional Model: fundamental aspects  When consulting the Cuban Encyclopedia of Ecured (Figure # 1) on the Professional Model in the Agricultural Engineering career, it is suggested that the graduate of these studies is a professional who is oriented to planning, managing, designing, evaluating and supervising the projects of engineering, aimed both at promoting social and productive development of the agricultural, forestry, livestock, agro-industrial and energy sectors, as well as at solving problems of conservation and use of natural resources and the environment. |



**Figura 1:** Aspectos fundamentales del Modelo de Profesional de Ingeniería Agrícola.

**Figure 1:** Fundamental aspects of the Agricultural Engineering Professional Model.

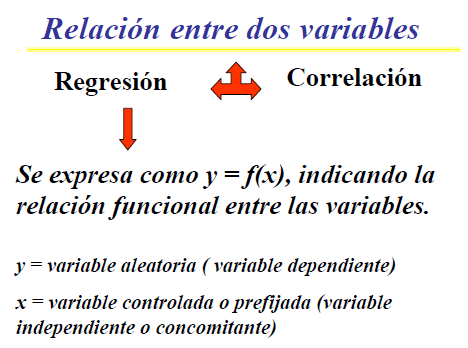
|  |  |
| --- | --- |
| Sobre la base del **Modelo de Profesional** a formal, es necesario la preparación de profesionales:   1. de perfil amplio; 2. con capacidades y métodos de ingeniería para el trabajo, la innovación tecnológica y la creatividad; 3. con adaptación rápida a tecnologías cambiantes en las condiciones de Cuba; 4. con capacidad para la acción y la educación social; 5. con los recursos humanos y materiales que se desarrollen en las universidades cubanas.   Al ampliarse el perfil ocupacional, la nueva carrera se llama Ingeniería Agrícola con los elementos siguientes:   * El **objeto de trabajo** es: Los sistemas de ingeniería agrícola para los procesos tecnológicos y biotecnológicos de la producción agropecuaria sostenible en los eslabones de base. * El **objeto de la profesión** es: La explotación de los sistemas de ingeniería agrícola para los procesos tecnológicos y biotecnológicos de la producción agropecuaria sostenible en los eslabones de base.   El principal problema profesionala resolver está relacionado con explotación de los sistemas de ingeniería para los procesos tecnológicos y biotecnológicos de la producción agropecuaria sostenible en sus eslabones de base.  Los sistemas de ingeniería agrícola se definen como la interacción multidisciplinaria y coherente de los principales conjuntos de elementos de las diferentes técnicas y tecnologías modernas que garantizan la producción agropecuaria sostenible, entre los que se incluyen:   1. La maquinaria agropecuaria, eléctrica, de tracción animal, para las casas de cultivo y las diferentes fuentes convencionales y alternativas de energía; 2. Los sistemas de riego, drenaje, desmonte, movimiento de tierra, transporte y casas de cultivo; 3. Los principales sistemas de post-cosecha; 4. Los equipos e instalaciones móviles y estacionarios para la producción agrícola, pecuaria, forestal, agroindustria rural y talleres de mantenimiento y reparación.   Los principales **problemas profesionales** que se presentan en la producción agropecuaria son:  • ¿Cómo administrar los sistemas de ingeniería para los procesos tecnológicos y biotecnológicos de la producción agropecuaria sostenible en sus eslabones de base?  • ¿Cómo perfeccionar los principales elementos de los sistemas de ingeniería para los procesos tecnológicos y biotecnológicos de la producción agropecuaria sostenible en sus eslabones de base?  • ¿Cómo mantener la capacidad de trabajo de los sistemas de ingeniería agrícola para los procesos tecnológicos y biotecnológicos de la producción agropecuaria sostenible en sus eslabones de base?  • ¿Cómo ejercer la docencia en el sistema de educación media y superior?  Los **campos de acción**, donde se manifiestan los métodos de esta ingeniería, son:  • La administración de los sistemas de ingeniería agrícola para los procesos tecnológicos y biotecnológicos de la producción agropecuaria sostenible.  • El perfeccionamiento de los principales elementos de los sistemas de ingeniería para los procesos tecnológicos y biotecnológicos de la producción agropecuaria sostenible en sus eslabones de base.  • El mantenimiento de la capacidad de trabajo de los sistemas de ingeniería agrícola para los procesos tecnológicos y biotecnológicos de la producción agropecuaria sostenible.  • El ejercicio de la docencia en el sistema de educación media y superior.  El **modo de actuación** del profesional es: Administrar, perfeccionar y mantener los sistemas de ingeniería agrícola para los procesos tecnológicos y biotecnológicos de la producción agropecuaria sostenible y ejercer la docencia en el sistema de educación media y superior.  Sus **esferas de actuación** son: los eslabones de base, donde se desarrollan los procesos tecnológicos y biotecnológicos de la producción agropecuaria sostenible para: los cultivos agrícolas; las especies animales, los procesos de poscosecha en la pequeña agroindustria rural, los servicios técnicos de ingeniería y en la docencia del sistema de educación media y superior.  Las principales **funciones profesionales** que debe desarrollar el egresado de esta carrera en los eslabones de base de la producción agropecuaria sostenible, acorde a los criterios de los empleadores y según la nomenclatura por ellos empleada, son:   * La administración (planificación, organización, dirección, control y operación) de los sistemas de ingeniería agrícola para la producción agropecuaria de los diferentes cultivos, especies animales, la poscosecha, el riego y el drenaje, mediante:   • La planificación y organización de los sistemas de ingeniería en los eslabones de base de la producción.  • La organización y operación de de los sistemas de ingeniería para las instalaciones pecuarias, el riego, el drenaje, el desmonte, la producción forestal, el transporte agrícola, los procesos de poscosecha y la industria rural.  • La ejecución y control de las programaciones para la utilización de la maquinaria, los sistemas de riego y la tracción animal en los sistemas de ingeniería agrícola con alta productividad y bajos costos.  • La determinación y evaluación de los índices técnico-económicos en la operación del parque de maquinaria, incluyendo los medios de tracción animal.  • La determinación y el control de la correcta agregación de los tractores, animales de trabajo, máquinas y aperos agrícolas y la regulación de los mismos, así como el correcto casamiento bomba-motor.  • La aplicación de los sistemas económicos y de administración para los procesos tecnológicos y biotecnológicos en los sistemas de ingeniería agrícola de la producción agropecuaria, así como en los talleres y demás instalaciones agropecuarias.  • La aplicación y perfeccionamiento de las normas de trabajo en las labores mecanizadas, de mantenimiento y reparación de la maquinaria, incluyendo los medios de tracción animal, así como controlando y evaluando la calidad del trabajo de los operadores de los medios motorizados, boyeros, obreros manuales y trabajadores de riego y de los talleres.  • La aplicación de las técnicas de seguridad y las normas de protección e higiene del trabajo.  • La interpretación y análisis de los resultados del trabajo de la maquinaria agropecuaria, incluyendo los sistemas de riego y drenaje, de los medios de tracción animal y la toma de decisiones que contribuyen a mejorar la disciplina tecnológica y laboral.   * El perfeccionamiento de los sistemas de ingeniería agrícola para los procesos tecnológicos y biotecnológicos de la producción agropecuaria sostenible en sus eslabones de base, mediante:   • La innovación y transferencia de nuevas tecnologías en los sistemas de ingeniería agrícola para los procesos tecnológicos y biotecnológicos de la producción agropecuaria.  • La confección de las cartas tecnológicas para la mecanización de los cultivos, incluyendo los medios de tracción animal y las producciones animal y forestal.  • La selección de los sistemas de ingeniería agrícola más convenientes para la ejecución de los distintos procesos, labores y operaciones productivos agropecuarios y forestales.  • El control y la evaluación de la calidad de las labores realizadas con los sistemas de ingeniería agrícola, según las exigencias técnicas de los distintos cultivos y explotaciones de producción animal y forestal.  • El análisis de los problemas existentes para el empleo de los sistemas de ingeniería agrícola en las distintas tecnologías agropecuarias, aportando soluciones a los procesos no resueltos.   * La explotación técnica de la maquinaria agropecuaria, incluyendo los medios motorizados, manuales y de tracción animal, los equipos de riego y de los medios de transporte, mediante:   • La confección, dirección y control de los planes de asistencia técnica a la maquinaria agropecuaria, incluyendo los equipos de riego y medios de transporte, así como para la restauración y fabricación de piezas de repuesto.  • La determinación del estado técnico de los equipos, máquinas e instalaciones agropecuarias, mediante su diagnóstico y evaluación.  • El desarrollo, modernización y perfeccionamiento de la base tecnológica de los talleres y equipos móviles de mantenimiento de acuerdo a las técnicas más avanzadas, así como la organización y control de los procesos tecnológicos (mantenimiento, conservación y reparación) que en los mismos se realizan.  • La planificación y el control de los insumos que requiere la explotación de los sistemas de ingeniería agrícola, que incluyen los medios motorizados, manuales y de tracción animal, equipos de riego y transporte, combustibles, lubricantes, piezas de repuesto, arreos y otros.   * La selección y uso de los distintos tipos alternativos de energía, mediante:   • La selección y el montaje de paneles solares en las granjas pecuarias.  • Selección, confección y puesta en marcha de plantas de biogás.  • Desarrollo y aplicación de la energía eólica para la extracción de agua y generación de corriente.  • El empleo de la biomasa como generador de energía y la cocción de alimentos.  • El estudio de los biocombustibles a partir de los aceites de residuo.  Los estudiantes del 3er año académico tienen como objetivo general:   * Interpretar como futuro ingeniero el papel activo y creador en el desarrollo de la producción agropecuaria sostenible, desde una óptica marxista, tercermundista, valorando las soluciones tecnológicas que brinda a los problemas profesionales, básicos, más frecuentes y generales, mediante sus dimensiones ética, sociocultural y humanística universal, cubana y de la profesión, de su impacto social y económico y con un enfoque dialéctico materialista. * Diseñar los órganos de trabajo y elementos principales de la maquinaria empleada en los sistemas de ingeniería agrícola -incluyendo para el riego, la tracción animal y el uso de energías renovables, exponiendo sus resultados de forma escrita y oral ante un tribunal y utilizando las estrategias curriculares.   Regresión lineal simple y Modelo de Análisis de Varianza simple y doble  Coincidimos con Fernández, Traparo y Domínguez (2010) al plantear de que existen magníficos libros sobre Diseños Experimentales y Análisis de Datos, pero las experiencias de los docentes investigadores del colectivo de Estadística Matemática de la UNAH, son necesarias para adentrar a los estudiantes de la Facultad de Ciencias Técnicas en el mundo de la experimentación y el correcto uso de una técnica en el procesamiento de la información.  Durante el curso de Estadística que reciben los estudiantes en 2do año de la carrera se han realizado análisis de los datos univariantes, donde observa en cada unidad elemental la población o de la muestra. A partir de este tema, se estudiarán poblaciones bivariantes, la cual es aquella en que se observan dos características en cada unidad elemental. Es lógico que pudieran ser observados más de dos características, convirtiéndose entonces el análisis en multivariante.  El análisis de más de una característica en estudio consiste en descubrir o medir la asociación o covariación entre las variables y determinar cómo ambas varían juntas. Si se puede establecer una relación media de alguna manera funcional en forma matemática, estaremos en disposición de estimar en forma precisa en promedio. Tal procedimiento se llama estimación por asociación.  Esta técnica de estimación por asociación es compleja y útil. Es un procedimiento de predicción, siendo la predicción una cuestión de gran importancia para la ciencia. Cualquier estudio científico se traza entre sus tareas principales, la de descubrir las relaciones generales entre las variables observadas y expresar la naturaleza de estas relaciones en forma matemáticamente precisa, de modo que se pueda predecir el valor una variable con base a otra.  Según criterios de Kuehl (2001) nadie ha tenido tanto impacto en los principios de la Estadística como Ronald A. Fisher, pero es el estadístico matemático Francis Galton quien introdujo el término de regresión en sus estudios de asociación. (Figura #2). | On the basis of the **Professional** **Model** to Formal, it is necessary to prepare professionals:   1. broad profile; 2. with capacities and engineering methods for work, technological innovation and creativity; 3. with rapid adaptation to changing technologies in Cuban conditions; 4. with capacity for action and social education; 5. with the human and material resources that are developed in Cuban universities.   As the occupational profile expands, the new career is called Agricultural Engineering with the following elements:  • The **object of work** is: Agricultural engineering systems for technological and biotechnological processes of sustainable agricultural production in the basic links.  • The **object of the profession** is: The exploitation of agricultural engineering systems for technological and biotechnological processes of sustainable agricultural production in the base links.  The main professional problem to be solved is related to the exploitation of engineering systems for technological and biotechnological processes of sustainable agricultural production in their base links.  Agricultural engineering systems are defined as the multidisciplinary and coherent interaction of the main sets of elements of the different modern techniques and technologies that guarantee sustainable agricultural production, which include:   1. Agricultural machinery, electrical, animal traction, for growing houses and different conventional and alternative sources of energy; 2. Irrigation, drainage, clearing, earthmoving, transportation and growing houses systems; 3. The main post-harvest systems; 4. Mobile and stationary equipment and facilities for agricultural, livestock, forestry, rural agroindustry production and maintenance and repair workshops.   The main **professional problems** that arise in agricultural production are:  • How to manage the engineering systems for the technological and biotechnological processes of sustainable agricultural production in their base links?  • How to improve the main elements of the engineering systems for the technological and biotechnological processes of sustainable agricultural production in their basic links?  • How to maintain the working capacity of agricultural engineering systems for technological and biotechnological processes of sustainable agricultural production in their base links?  • How to teach in the secondary and higher education system?  The **fields of action**, where the methods of this engineering are manifested, are:  • The administration of agricultural engineering systems for the technological and biotechnological processes of sustainable agricultural production.  • The improvement of the main elements of the engineering systems for the technological and biotechnological processes of sustainable agricultural production in their basic links.  • Maintaining the work capacity of agricultural engineering systems for technological and biotechnological processes of sustainable agricultural production.  • The exercise of teaching in the secondary and higher education system.  The professional's **mode of action** is: Manage, improve and maintain agricultural engineering systems for technological and biotechnological processes of sustainable agricultural production and teach in the secondary and higher education system.  His **spheres of action** are: the base links, where the technological and biotechnological processes of sustainable agricultural production are developed for: agricultural crops; animal species, post-harvest processes in small rural agribusiness, technical engineering services, and teaching in the secondary and higher education system.  The main professional functions that the graduate of this career must develop in the basic links of sustainable agricultural production, according to the criteria of the employers and according to the nomenclature used by them, are:   * The administration (planning, organization, direction, control and operation) of agricultural engineering systems for agricultural production of different crops, animal species, post-harvest, irrigation and drainage, through:   • The planning and organization of engineering systems in the base links of production.  • The organization and operation of engineering systems for livestock facilities, irrigation, drainage, clearing, forestry production, agricultural transportation, post-harvest processes and rural industry.  • The execution and control of schedules for the use of machinery, irrigation systems and animal traction in agricultural engineering systems with high productivity and low costs.  • The determination and evaluation of the technical-economic indices in the operation of the machinery park, including the means of animal traction.  • The determination and control of the correct aggregation of tractors, work animals, agricultural machines and implements and their regulation, as well as the correct pump-motor match.  • The application of economic and administration systems for technological and biotechnological processes in agricultural engineering systems of agricultural production, as well as in workshops and other agricultural facilities.  • The application and improvement of work standards in mechanized, maintenance and repair work of machinery, including animal traction means, as well as controlling and evaluating the quality of the work of operators of motorized means, cattlemen, workers manuals and irrigation and workshop workers.  • The application of safety techniques and standards of protection and hygiene at work.  • The interpretation and analysis of the results of the work of the agricultural machinery, including the irrigation and drainage systems, the means of animal traction and the decision-making that contributes to improving the technological and labor discipline.   * The improvement of agricultural engineering systems for the technological and biotechnological processes of sustainable agricultural production in their basic links, through:   • Innovation and transfer of new technologies in agricultural engineering systems for technological and biotechnological processes in agricultural production.  • The preparation of technological charts for the mechanization of crops, including means of animal traction and animal and forestry productions.  • The selection of the most convenient agricultural engineering systems for the execution of the different agricultural and forestry productive processes, tasks and operations.  • The control and evaluation of the quality of the work carried out with the agricultural engineering systems, according to the technical requirements of the different crops and animal and forestry production farms.  • The analysis of the existing problems for the use of agricultural engineering systems in different agricultural technologies, providing solutions to unresolved processes.   * The technical exploitation of agricultural machinery, including motorized, manual and animal-drawn means, irrigation equipment and means of transport, through:   • The preparation, management and control of technical assistance plans for agricultural machinery, including irrigation equipment and means of transport, as well as for the restoration and manufacture of spare parts.  • Determination of the technical state of agricultural equipment, machines and facilities, through diagnosis and evaluation.  • The development, modernization and improvement of the technological base of the workshops and mobile maintenance teams according to the most advanced techniques, as well as the organization and control of the technological processes (maintenance, conservation and repair) that are carried out in them .  • The planning and control of the inputs required for the exploitation of agricultural engineering systems, which include motorized, manual and animal traction means, irrigation and transport equipment, fuels, lubricants, spare parts, harnesses and others.   * The selection and use of the different alternative types of energy, through:   • The selection and assembly of solar panels in livestock farms.  • Selection, preparation and start-up of biogas plants.  • Development and application of wind energy for the extraction of water and power generation.  • The use of biomass as an energy generator and food cooking.  • The study of biofuels from waste oils.  The students of the 3rd academic year have as a general objective:   * Interpret as a future engineer the active and creative role in the development of sustainable agricultural production, from a Marxist, Third World perspective, valuing the technological solutions that it provides to professional, basic, more frequent and general problems, through its ethical, sociocultural dimensions and universal humanistic, Cuban and of the profession, of its social and economic impact and with a materialistic dialectical approach. * Design the working organs and main elements of the machinery used in agricultural engineering systems -including for irrigation, animal traction and the use of renewable energies, presenting their results in writing and orally before a court and using the strategies curricular.   Simple Linear Regression and Simple and Double Analysis of Variance Model  We agree with Fernández, Traparo and Domínguez (2010) in stating that there are magnificent books on Experimental Designs and Data Analysis, but the experiences of the research professors of the UNAH group of Mathematical Statistics are necessary to introduce the students of the Faculty of Technical Sciences in the world of experimentation and the correct use of a technique in information processing.  During the Statistics course received by students in the 2nd year of the degree, univariate data analysis has been carried out, where the population or sample is observed in each elementary unit. From this topic, bivariate populations will be studied, which is one in which two characteristics are observed in each elementary unit. It is logical that more than two characteristics could be observed, thus making the analysis multivariate.  The analysis of more than one characteristic under study consists of discovering or measuring the association or covariation between the variables and determining how both vary together. If an average relationship can be established in some functional way mathematically, we will be in a position to accurately estimate the average. Such a procedure is called estimation by association.  This estimation technique by association is complex and useful. It is a prediction procedure, prediction being a matter of great importance to science. Any scientific study is drawn among its main tasks that of discovering the general relationships between the observed variables and expressing the nature of these relationships mathematically precisely, so that the value of one variable can be predicted based on another.  According to Kuehl's (2001) criteria, no one has had as much impact on the principles of Statistics as Ronald A. Fisher, but it is the mathematical statistician Francis Galton who introduced the term of regression in his association studies. (Figure # 2). |

****

**Figura 2**: El término regresión según Francis Galton.

**Figure 2:** The term regression according to Francis Galton.

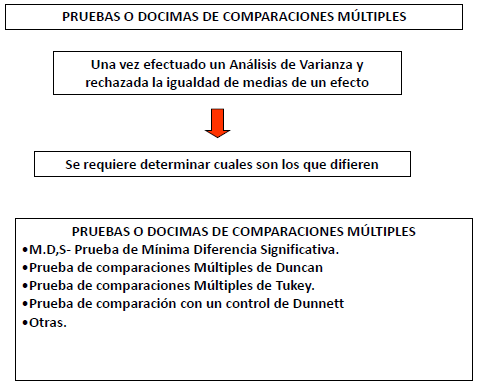
|  |  |
| --- | --- |
| Existen dos aspectos distintos, pero relacionados, del estudio de la asociación entre variables. Un primer aspecto del análisis de asociación se conoce como **análisis de correlación** Guerra *et al*. (2006), el cual se ocupa de determinar ¨el grado de relación¨ entre las variables, en este caso las de designación de las variables dependiente e independiente es una elección estrictamente personal y no tiene significación práctica.  El segundo aspecto se llama **análisis de regresión** Guerra *et al*. (2006) que trata de establecer la forma de la relación entre las variables; es decir, en el análisis de regresión se estudia la relación funcional entre las variables, de modo que podemos predecir el valor de una con base en otra u otras. Convenientemente la variable o variables que son la base de la predicción se llaman variable o variables independientes y la variable que se va a predecir se denomina variable dependiente. (Figura #3) | There are two different but related aspects of the study of the association between variables. A first aspect of the association analysis is known as the **correlation analysis** Guerra et al. (2006), which deals with determining ¨the degree of relationship¨ between the variables, in this case the designation of the dependent and independent variables is a strictly personal choice and has no practical significance.  The second aspect is called **regression analysis** Guerra et al. (2006) that tries to establish the form of the relationship between the variables; that is, in the regression analysis the functional relationship between the variables is studied, so that we can predict the value of one based on another or others. Conveniently the variable or variables that are the basis of the prediction are called the independent variable or variables and the variable to be predicted is called the dependent variable. (Figure # 3) |

****

**Figura 3**: El método de regresión.

**Figure 3**: The regression method.

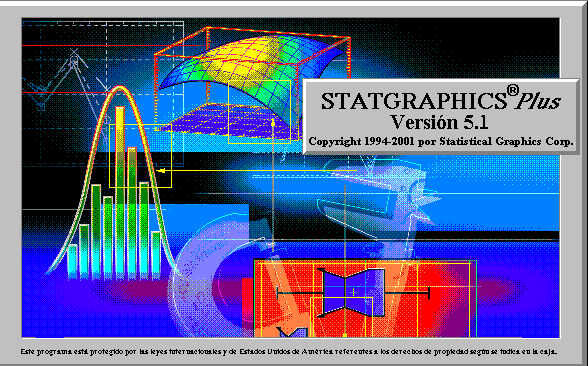
|  |  |
| --- | --- |
| Según Hernández *et al*. (2018) la regresión lineal es un modelo matemático para estimar el efecto de una variable sobre otra y que está asociado con el coeficiente r de Pearson, cuya base es el diagrama de dispersión.  El modelo lineal de regresión simple Guerra *et al*. (2006) se usa para modelar un fenómeno donde el comportamiento de una variable aleatoria es explicada por una o más variables explicativas. La regresión se expresa como  indicando la relación funcional entre variables, donde  es la variable dependiente (variable aleatoria) y  es la variable independiente o concomitante (variable controlada o prefijada).  Se pueden formar pares ordenados para representar en un sistema de coordenadas cartesianas, a la cual se denomina diagrama de dispersión. El comportamiento de estos diagramas pueden ser lineales o no lineales.  Cada observación pudiera expresarse como donde:  variable en estudio.  constantes que se denominan coeficientes de regresión.  error aleatorio o perturbaciones.  El coeficiente de determinación (R2) es un indicador descriptivo que expresa el porciento de la variación total que es explicada por este modelo de regresión:    En:  Si  hay una ausencia total de ajuste en el modelo.  Si  es bastante poco probable, ya que debe existir algún margen de error.  Por ejemplo, si se desea analizar el precio de una pieza de repuesto para combinada cañera, hay razones para pensar que dicho precio, puede tener relación con la cantidad de piezas producidas. Haciendo abstracción de otras variables y posibles relaciones que puedan estar presentes, se pudiera cuestionar si el precio de una pieza de repuesto, podrá expresarse en función de la demanda de piezas a través de una relación lineal simple.  El conocimiento de esta técnica estadística por parte de cualquier investigador le proporciona una herramienta muy útil y poderosa para el trabajo investigativo. Esta técnica estadística fue desarrollada por R. A. Fisher. El análisis de varianza permite el estudio de las características medidas u observadas, cuyos valores dependan de varias clases de efectos que operan simultáneamente, y mediante este análisis poder decidir si tales efectos son o no diferentes.   * ANOVA simple (modelo lineal):   Los matemáticos estadísticos Hernández, Fernández y Batista (2018) plantean que el análisis de varianza (ANOVA) es una prueba estadística para analizar si dos o más grupos difieren significativamente entre sí, en cuanto a sus medias y varianzas.  Bouza y Sistachs (2006) plantean que cuando tratamos el estudio de varias poblaciones, la comparación de las medias es llevada a cabo utilizando un modelo: el modelo lineal.  En este modelo se asume que el valor de la variable en la respuesta obtenida en la población i en la unidad j es descrito por un factor asociado a la población general, y por un efecto asociado a la subpoblación correspondiente.  La clasificación simple se describe como Yij = µ + αi + εij y debe cumplir un grupo de requerimientos a cumplir.  Aun cuando la normalidad pueda ser falsa, la robustez de los métodos del análisis de varianza permite utilizar, sin que haya grandes afectaciones de la potencia y el nivel de significación de la prueba, los procedimientos que se presentan en esta prueba o dócima de hipótesis.  Este modelo está justificado para diseños completamente aleatorizados, donde las unidades experimentales son homogéneas (estar seguros que la única causa de la variación es ocasionada por los tratamientos en estudio, aparte del error experimental).  En la Figura #4 se grafican aspectos de contenidos muy importantes sobre las pruebas o dócimas de comparación múltiples que son tratadas por los estudiantes en la asignatura Estadística que reciben en segundo año de la carrera. | According to Hernández et al. (2018) linear regression is a mathematical model to estimate the effect of one variable on another and that is associated with Pearson's r coefficient, whose base is the scatter diagram.  The simple linear regression model Guerra et al. (2006) is used to model a phenomenon where the behavior of a random variable is explained by one or more explanatory variables. Regression is expressed as indicating the functional relationship between variables, where is the dependent variable (random variable) and is the independent or concomitant variable (controlled or prefixed variable).  Ordered pairs can be formed to represent in a Cartesian coordinate system, which is called a scatter diagram. The behavior of these diagrams can be linear or non-linear.  Each observation could be expressed as where:  variable under study.  constants called regression coefficients.  random error or disturbances.  The coefficient of determination (R2) is a descriptive indicator that expresses the percentage of the total variation that is explained by this regression model:      In:  If there is a total absence of fit in the model.  If it is quite unlikely, as there must be some margin of error.  For example, if you want to analyze the price of a spare part for a sugarcane combination, there are reasons to think that this price may be related to the quantity of parts produced. Making abstraction of other variables and possible relationships that may be present, it could be questioned whether the price of a spare part can be expressed as a function of the demand for parts through a simple linear relationship.  The knowledge of this statistical technique by any researcher provides them with a very useful and powerful tool for investigative work. This statistical technique was developed by R. A. Fisher. The analysis of variance allows the study of the measured or observed characteristics, whose values ​​depend on several classes of effects that operate simultaneously, and by means of this analysis to be able to decide whether or not such effects are different.  • Simple ANOVA (linear model):  Statistical mathematicians Hernández, Fernández and Batista (2018) suggest that the analysis of variance (ANOVA) is a statistical test to analyze whether two or more groups differ significantly from each other, in terms of their means and variances.  Bouza and Sistachs (2006) state that when we deal with the study of several populations, the comparison of means is carried out using a model: the linear model.  In this model it is assumed that the value of the variable in the response obtained in population i in unit j is described by a factor associated with the general population, and by an effect associated with the corresponding subpopulation.  The simple classification is described as Yij = µ + αi + εij and must meet a group of requirements to meet.  Even when normality may be false, the robustness of the analysis of variance methods allows the use of the procedures presented in this test or hypothesis test without major effects on the power and significance level of the test.  This model is justified for completely randomized designs, where the experimental units are homogeneous (be sure that the only cause of the variation is caused by the treatments under study, apart from the experimental error).  In Figure # 4, very important content aspects are graphed on the tests or multiple comparison tests that are treated by the students in the Statistics subject that they receive in the second year of the degree. |



**Figura 4**: Sobre las pruebas de comparación múltiples.

**Figure 4**: About multiple comparison tests.

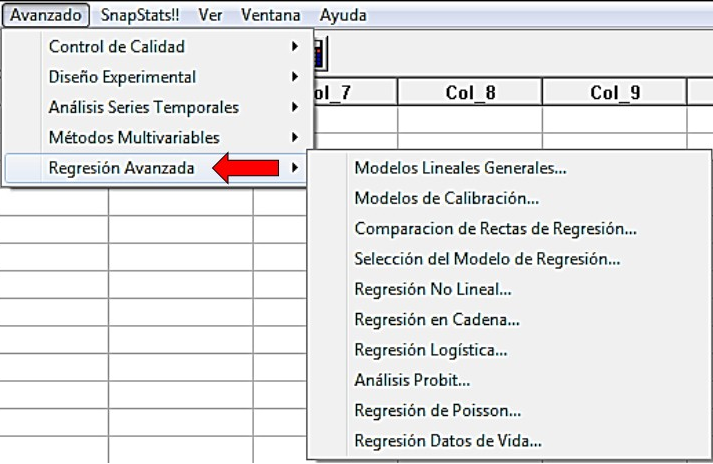
|  |  |
| --- | --- |
| La prueba Tukey brinda grandes facilidades para las comparaciones múltiples por la protecciones ante los errores que se presentan (Tukey, 1977).  Según estudios realizados por Di Rienzo *et al*. (2005) se señala que la prueba de Tukey es muy efectiva ya que examina con un solo estadístico todas las diferencias de medias muestrales en estudio, pero cuando los tamaños de muestra son muy diferentes, esta prueba deja de ser confiable, caso en el cual puede utilizarse la prueba de Scheffé (no considerada en la tabla anterior), pero que el paquete estadístico si la determina.   * ANOVA doble (modelo lineal).   En este modelo las unidades experimentales no son homogéneas, y se permite controlar el efecto de la no homogeneidad de tal forma que el error experimental se disminuya.  Por ejemplo, si se quiere conocer si el grado de enyerbamiento promedio (en kg de masa verde/m2) de un suelo varía cuando se aplican diferentes tecnologías de preparación de suelo. Se cuenta con área de cultivo homogénea, al igual que el resto de las condiciones agrícolas. En esta situación debe cuestionarse si tiene sentido pensar en un Análisis de regresión o en un Análisis de Varianza.  La clasificación doble se describe como Yij = µ + αi + βj + εij y debe cumplir un grupo de requerimientos a cumplir.  Ante la disyuntiva de cuántas variables independientes e dependientes deben incluirse en un experimento . Hernández *et al*. (2018) plantean que no hay reglas para ello, depende de cómo haya sido planteado el problema de la investigación y las limitaciones que existan.  Modelo lineal general. Hipótesis de base del modelo lineal y cumplimiento de las mismas. Clasificación del modelo lineal atendiendo a los Xji y los βj: modelo de regresión y modelo de análisis de varianza  Los estudios realizados por Cué *et al*. (1987) señalan que la estimación de los parámetros de un modelo particular de regresión es equivalente a estimar el vector β de su correspondiente forma matricial. Los modelos de regresión son funciones de regresión lineal.   * Supuestos de Hipótesis de base del modelo lineal.   En la formulación del modelo lineal se hace necesaria la introducción de un conjunto de suposiciones que por la importancia que encierran en los diferentes pasos del proceso de formulación, estimación y análisis del modelo se las denomina hipótesis de base; sin su observancia carecería de sentido la modelación de la situación.   * Aditividad de efectos: Suposición que está implícita en el modelo desde el momento en que se asume que la variable aleatoria y puede expresarse como una combinación lineal de p efectos βj. * Distribución Normal de los errores: Se asume que los ei son variables aleatorias normalmente distribuidas, o sea . * Incorrelación de errores: Los ei son variables aleatorias independientes entre sí. * Homogeniedad de varianza: Los ei son variables aleatorias con media cero y varianza homogénea.   Lo planteado anteriormente es equivalente para la variable respuesta o variable dependiente  Yi, que también distribuyen Normal e independientes entre sí.  Clasificación de modelo atendiendo a la naturaleza de los x ji.   * Modelo de análisis de regresión:   Se dice que el modelo lineal es un modelo de análisis de regresión cuando los valores xji provienen de variables continuas o discretas no contadoras.   * Modelo de Análisis de Varianza:   Se dice que el Modelo Lineal General es un modelo de Análisis de Varianza cuando los valores de los xji son provenientes de variables indicadoras o contadoras cuyos valores son 0 o 1 e indican presencia o ausencia del bj correspondiente en una observación dada. Se retomará de nuevo el Modelo Lineal General de ANOVA cuando se analice el Tema 2 de la asignatura sobre los Diseños de Experimentales.  Modelo Lineal General de Regresión para dos o más variables    Donde i=1,2 …, n y j=1, 2, …., p  donde Yi es la variable respuesta o variable dependiente (también se le denomina variable explicada, endógena, regresante o aleatoria), Xpi es la variable independiente, explicatoria o explicativa (también se le denomina variable exógena, regresora, concomitante, o variable continua o discreta no contadora), X0i=1 y β0 es el término independiente, βj es el coeficiente de regresión parcial (mide el efecto marginal sobre la variable respuesta por unidad de variación de una de las variable Xji cuando las otras variables pertenecen constantes) y Ei son las perturbaciones aleatorias o errores aleatorios, los cuales se distribuyen Normal con media cero y varianza constante, es decir que Ei (0, σ2).  Son Modelos que cumplen los mismos supuestos teóricos, además de la no multicolinealidad, es decir, las variables independientes tienen que ser linealmente independiente, pero lo veremos en detalle posteriormente.  Por ejemplo, si se desea analizar el precio de una pieza de repuesto para combinada cañera, y la posible relación con la cantidad de dichas piezas, pudiera pensarse que también se relacione con el precio de su materia prima fundamental. En este nuevo contexto se puede cuestionar cómo podrá expresarse la relación funcional.  El modelo que debe seguirse sería establecido como:    Donde i=1,2 …, n y j=1, 2, …., p  donde la variable yi es el precio de la pieza de repuesto i (variable a explicar), x1j es la cantidad de piezas de repuesto ij (variable explicativa), v2j es el precio de su materia prima (variable explicativa), β0, β1, β2 son constantes, denominadas coeficientes de regresión parciales, y eij son los errores experimentales.  Uso del paquete estadístico Statgraphics plus 5.1  Los paquetes estadísticos han venido a facilitar el trabajo en el procesamiento de datos para un investigador. Son múltiples y muy variados, pero el Statgraphics plus 5.1 (Figura#5) ha sido asumido por el colectivo de docentes e investigadores de la UNAH, para el pregrado en diferentes carreras de la Universidad. Es por ello que la incorporación de esta herramienta, en la docencia de los estudiantes debe hacer de modo gradual, destacando no solo sus posibilidades de análisis, simulación y visualización, sino sobre todo, enfatizando la interpretación adecuada de los resultados y su puesta en relación con las preguntas que guían el análisis de los datos. | The Tukey test provides great facilities for multiple comparisons due to the protections against errors that occur (Tukey, 1977).  According to studies by Di Rienzo et al. (2005) point out that Tukey's test is very effective since it examines all the differences in the sample means under study with a single statistic, but when the sample sizes are very different, this test is no longer reliable, in which case The Scheffé test (not considered in the previous table) can be used, but the statistical package does determine it.  • Double ANOVA (linear model).  In this model, the experimental units are not homogeneous, and it is allowed to control the effect of inhomogeneity in such a way that the experimental error is reduced.  For example, if you want to know if the average degree of mulching (in kg of green mass / m2) of a soil varies when different soil preparation technologies are applied. There is a homogeneous growing area, like the rest of the agricultural conditions. In this situation it must be questioned whether it makes sense to think of a Regression Analysis or an Analysis of Variance.  The double classification is described as Yij = µ + αi + βj + εij and must meet a group of requirements to meet.  Faced with the dilemma of how many independent and dependent variables should be included in an experiment. Hernández et al. (2018) state that there are no rules for this, it depends on how the research problem has been raised and the limitations that exist.  General linear model. Basic hypothesis of the linear model and compliance with them. Classification of the linear model according to the Xji and βj: regression model and analysis of variance model  The studies carried out by Cué et al. (1987) point out that the estimation of the parameters of a particular regression model is equivalent to estimating the vector β of its corresponding matrix form. Regression models are linear regression functions.  • Assumptions of basic hypotheses of the linear model.  In the formulation of the linear model, it is necessary to introduce a set of assumptions that, due to the importance they hold in the different steps of the process of formulation, estimation and analysis of the model, are called the base hypothesis; without its observance, the modeling of the situation would be meaningless.   * Additivity of effects: Assumption that is implicit in the model from the moment it is assumed that the random variable y can be expressed as a linear combination of p effects βj. * Normal distribution of errors: It is assumed that the ei are normally distributed random variables, that is. * Correlation of errors: The ei are independent random variables. * Homogeneity of variance: The ei are random variables with zero mean and homogeneous variance.   The above is equivalent for the response variable or dependent variable  Yi, which also distribute Normal and independent of each other.  Classification of the model according to the nature of the x ji.   * Regression analysis model:   The linear model is said to be a regression analysis model when the xji values ​​come from non-counting discrete or continuous variables.   * Analysis of Variance Model:   It is said that the General Linear Model is an Analysis of Variance model when the values ​​of the xji are coming from indicator or counting variables whose values ​​are 0 or 1 and indicate the presence or absence of the corresponding bj in a given observation. The ANOVA General Linear Model will be taken up again when Topic 2 of the subject on Experimental Designs is analyzed.  General Linear Regression Model for two or more variables    Donde i=1,2 …, n y j=1, 2, …., p  where Yi is the response variable or dependent variable (it is also called the explained, endogenous, regressive or random variable), Xpi is the independent, explanatory or explanatory variable (it is also called the exogenous, regressive, concomitant variable, or the continuous or discrete variable non-counting), X0i = 1 and β0 is the independent term, βj is the partial regression coefficient (measures the marginal effect on the response variable per unit of variation of one of the variables Xji when the other variables belong to constants) and Ei are the random disturbances or random errors, which are Normally distributed with zero mean and constant variance, that is, Ei (0, σ2).  They are Models that meet the same theoretical assumptions, in addition to non-multicollinearity, that is, the independent variables have to be linearly independent, but we will see it in detail later.  For example, if you want to analyze the price of a spare part for a sugarcane combination, and the possible relationship with the quantity of these parts, it could be thought that it is also related to the price of its fundamental raw material. In this new context it can be questioned how the functional relationship can be expressed.  The model to be followed would be established as:    Donde i=1,2 …, n y j=1, 2, …., p  where the variable yi is the price of the spare part i (variable to be explained), x1j is the quantity of spare parts ij (explanatory variable), v2j is the price of their raw material (explanatory variable), β0, β1, β2 are constants, called partial regression coefficients, and eij are the experimental errors.  Using the statistical package Statgraphics plus 5.1  Statistical packages have come to facilitate the work in data processing for a researcher. They are multiple and very varied, but Statgraphics plus 5.1 (Figure # 5) has been assumed by the UNAH group of teachers and researchers, for undergraduates in different University careers. That is why the incorporation of this tool in the teaching of students should be done gradually, highlighting not only its possibilities of analysis, simulation and visualization, but above all, emphasizing the adequate interpretation of the results and their relationship with the questions that guide the data analysis. |

****

**Figura 5**: Presentación de inicio del Statgraphics plus 5.1.

**Figure 5**: Statgraphics plus 5.1 startup presentation.

|  |  |
| --- | --- |
| El Statgraphics plus 5.1 es un paquete estadístico profesional que proporciona los tipos de análisis estadísticos comunes, y brinda otros instrumentos necesarios en el análisis de los datos, tales el editor de datos, utilidades para manejar los ficheros de datos, opciones para cambiar parámetros del sistema y ayuda. Además, utiliza varios tipos de ficheros como son:   1. Ficheros de datos: En ellos se introducen los datos a procesar. Para realizar un análisis estadístico es necesario que haya un fichero de datos abiertos. 2. Ficheros Statfolio: Son ficheros que se pueden grabar con los resultados del análisis, para tenerlos disponibles en el futuro. Este incluye también el Fichero de Datos, así como todas las ventanas de resultados que no se hubieran cerrado al grabar el Statfolio. 3. Statreporter: Sirve para escribir el informe a medida que se analizan los datos.   Se recomienda a los usuarios que al instalar esta versión, se deben asegurar de que esté conectado al sistema a través de un identificador con privilegios de administrador, los que le permitirán instalar los controladores centinela.  Debe disponer de al menos 65 MB de espacio libre en su disco duro para instalar los ficheros de programa, documentación y ficheros de datos de ejemplo. Se requiere un mínimo de 30 MB para la instalación mínima de sólo los ficheros de programa. Además, se solicita un mínimo de 32 MB de RAM con un rendimiento medio. Para cálculos y gráficos complejos será necesaria memoria RAM adicional. El disco de la versión 5.1 contiene 24 ficheros de instalación.  Este programa está estructurado mediante una serie de Menús encadenados, cada uno de los cuales tiene diversas opciones que se pueden usar para cambiar los resultados de los análisis o para pedir nuevos análisis. El funcionamiento de todos las acciones es muy parecido, de modo que todo lo que se aprende en un curso inicial servirá para avanzar posteriormente o incluso aprender el uso de otros programas, ya que la mayoría tienen una estructura y filosofía similar.  Batanero y Díaz (2015) señalaron que en el Statgraphics hay varias vías relacionadas con la regresión. Estos son:   * A partir de la opción del menú principal Dependencia, se obtiene una ventana con Regresión Simple… (Figura #6). La entrada de las variables pide la variable Y (variable dependiente o explicada) y la variable X (variable independiente o explicativa).     Figura #6: Opción Dependencia – Regresión Simple.   * A partir de la opción del menú principal Dependencia, se obtiene una ventana con Regresión Múltiple… (Figura #7). La entrada de las variables pide la variable Y (variable dependiente o explicada) y todas las variables X (variables independientes o explicativas).     Figura #7: Opción Dependencia – Regresión Múltiple.   * A partir de la opción del menú principal Avanzado, se obtiene una ventana con Regresión Avanzada… (Figura #8). La entrada de las variables pide la variable Y (variable dependiente o explicada) y todas las variables X (variables independientes o explicativas) como factores categóricos y factores cuantitativos. | Statgraphics plus 5.1 is a professional statistical package that provides the common types of statistical analysis, and provides other tools necessary for data analysis, such as the data editor, utilities to manage data files, options to change system parameters. and help. In addition, it uses several types of files such as:  1. Data files: The data to be processed are entered in them. To perform a statistical analysis it is necessary to have an open data file.  2. Statfolio files: These are files that can be recorded with the results of the analysis, to have them available in the future. This also includes the Data File, as well as all the results windows that would not have been closed when saving the Statfolio.  3. Statreporter: Used to write the report as the data is analyzed.  Users are advised to install this version to ensure that they are logged into the system through an identifier with administrator privileges, which will allow them to install the sentinel drivers.  You must have at least 65 MB of free space on your hard disk to install the program files, documentation, and sample data files. A minimum of 30 MB is required for the minimal installation of only the program files. Additionally, a minimum of 32MB of RAM is requested with average performance. For complex calculations and graphs additional RAM will be required. The version 5.1 disc contains 24 installation files.  This program is structured through a series of chained Menus, each of which has various options that can be used to change the results of the analyzes or to request new analyzes. The operation of all the actions is very similar, so that everything that is learned in an initial course will serve to advance later or even learn the use of other programs, since most have a similar structure and philosophy.  Batanero and Díaz (2015) pointed out that there are several pathways related to regression in Statgraphics. These are:   * From the main menu option Dependency, a window with Simple Regression ... is obtained (Figure # 6). The input of the variables asks for the variable Y (dependent or explained variable) and the variable X (independent or explanatory variable).       Figure # 6: Dependency Option - Simple Regression.   * A partir de la opción del menú principal Dependencia, se obtiene una ventana con Regresión Múltiple… (Figura #7). La entrada de las variables pide la variable Y (variable dependiente o explicada) y todas las variables X (variables independientes o explicativas).     Figure # 7: Option Dependency - Multiple Regression.   * From the Advanced main menu option, you get a window with Advanced Regression… (Figure # 8). The input of the variables asks for the variable Y (dependent or explained variable) and all the variables X (independent or explanatory variables) as categorical factors and quantitative factors. |



**Figura 8:** Opción Avanzado – Regresión Avanzada.

**Figure 8:** Advanced Option - Advanced Regression.

|  |  |
| --- | --- |
| Es importante destacar que darse cuenta de cuál es la variable que debe ser Y y cuál debe ser la variable X porque el programa encontrará una ecuación de Y en función de X (que no siempre coincide con la ecuación de X en función de Y). Que una variable quede explicada por otra no quiere decir que haya una relación de causa y efecto.  El Statgraphics tiene diversas representaciones gráficas. La más útil es la de Gráfico de Modelo Ajustado que dibuja la curva ajustada sobre la nube de puntos. Cambiando el tipo de función en Opciones de Análisis se puede observar visualmente cuál de los modelos es el más ajustado a los datos.  Vale la pena señalar que los estudios realizados por Freixa *et al*. (2018) sobre la exploración de la relación bivariable con línea ¨resistente¨ es una forma de relación entre variables. Es especialmente necesario cuando se pretende modelizar una relación mediante estrategia lineal y no se dispone de suficiente conocimiento previo de que tal relación es factible | It is important to note that realizing which variable should be Y and what variable X should be because the program will find an equation of Y as a function of X (which does not always coincide with the equation of X as a function of Y). That one variable is explained by another does not mean that there is a cause and effect relationship.  Statgraphics has different graphical representations. The most useful is the Fitted Model Graph which draws the fitted curve over the point cloud. By changing the type of function in Analysis Options, it can be visually observed which of the models is the most adjusted to the data.  It is worth noting that the studies conducted by Freixa et al. (2018) on the exploration of the bivariate relationship with a ¨resistant¨ line is a form of relationship between variables. It is especially necessary when trying to model a relationship using a linear strategy and there is not enough prior knowledge that such a relationship is feasible. |

Bibliografías / References

Batanero, C; Díaz, M. C. 2015.Análisis de datos con Statgraphics. Departamento de Didáctica de la Matemática. Universidad de Granada. Editorial La Gioconda S. L. España.

Bouza, C. N; Sistachs, V. 2006. Estadística. Teoría básica y ejercicios. Editorial Félix Varela. La Habana, Cuba.

Cué, J. L; Castell, E; Hernández, J. M. 1987.Estadística, 2da. Parte. Universidad de la Habana. Empresa Nacional de producción del MES ¨Andre Voisin¨. Ciudad de la Habana, Cuba.

Di Rienzo, J. A; Casanoves, F; González, L. A; Tablada, E. M; Díaz, M. del P; Robledo, C. W; Balzarini, M. G. 2005.Estadística para las Ciencias Agropecuarias. Sexta edición. Edición Electrónica. Córdova, Argentina.

Fernández, R; Traparo, A; Domínguez, J. 2010.Experimentación en Agricultura. Consejería de Agricultura y Pesca. Junta de Andalucía. Editorial Serie. Sevilla, España.

Freixa, M; Salafranca, I; Guardia, J; Ferrer, R; Turbany, J. 2018.Análisis exploratorio de datos: Nuevas técnicas estadísticas. Universidad de Barcelona. Editorial PPU, S.A. España.

Guerra, C. W; Menéndez, E; Barrero, R; Egaña, E. 2006.Estadística. La Habana: Editorial Pueblo y Educación. Cuba.

Hernández, R; Fernández C; Batista, P. 2018.Metodología de la Investigación. 3ra. Edición. Editorial Mc.Graw-Hill Interamericana. México.

Kuehl, R. O. 2001.Diseño de Experimentos: Principios estadísticos para el diseño y análisis de investigaciones. Segunda edición. Editorial Thomson Learning. México D. F.

Miller, I; Freund, J. E; Johnson, R. 2007. Probabilidades y estadísticas para ingenieros. Parte I. Editorial Félix Varela. La Habana, Cuba.

Montgomery, D. C. 1991. Diseño y análisis de experimentos. Grupo editorial Iberoamericana S. A. México D. F.

Tukey, J. W. 1977.Exploratory Data Analysis. Editorial Addison´Wesley. Massachussets, USA.

Walpole, R. E; Myers, R. M; Myers, S. L. 2008. Probabilidad y estadística para ingenieros. Parte II. Editorial Félix Varela. La Habana, Cuba.