GRUPO BLR

DISEÑO EXPERIMENTAL

INTRODUCCIÓN A LA EXPERIMENTACIÓN Y PROCESAMIENTO DE DATOS EN INVESTIGACIONES AGROPECUARIAS

ISBN: 978-9907-852-14-1

Irán Rodríguez Delgado
Hipólito Israel Pérez Iglesias

2026

DISEÑO EXPERIMENTAL INTRODUCCIÓN A LA EXPERIMENTACIÓN Y PROCESAMIENTO DE DATOS EN INVESTIGACIONES AGROPECUARIAS

Irán Rodríguez Delgado

Hipólito Israel Pérez Iglesias

AUTORES

UTMACH

Este libro ha sido debidamente examinado y valorado en la modalidad doble par ciego con fin de garantizar la calidad científica.

©Grupo Editorial BLR
Universidad Técnica de Machala
Riobamba – Ecuador
Correo: publicaciones@grupoblr.com
https://grupoblr.com/libros-investigo
REPOSITORIO



Rodríguez, I., Pérez., H. (2026) Diseño experimental introducción a la experimentación y procesamiento de datos en investigaciones agropecuarias. Grupo Editorial BLR.

© Irán Rodríguez Delgado
Hipólito Israel Pérez Iglesias

Diseño experimental introducción a la experimentación y procesamiento de datos en investigaciones agropecuarias.

ISBN: 978-9907-852-14-1

El copyright promueve la libertad de expresión, protege la diversidad de ideas y conocimiento, además apoya la libre expresión. Se prohíbe de manera rigurosa la producción o el almacenamiento de esta publicación, ya sea en su totalidad o en parte, está estrictamente prohibido por ley, incluyendo el diseño de la portada, así como su difusión a través de cualquiera de sus medios, ya sean electrónicos, mecánicos, ópticos, de grabación o incluso de fotocopia, sin permiso de los propietarios de los derechos de autor.

Filiación de autores

Irán Rodríguez Delgado

Universidad Técnica de Machala

Correo Electrónico: irodriguez@utmachala.edu.ec

ORCID: https://orcid.org/0000-0002-6453-2108

Hipólito Israel Pérez Iglesias

Universidad Técnica de Machala

Correo Electrónico: hipolitoperez1944@gmail.com

ORCID: https://orcid.org/0000-0002-3368-8716

PREFACIO

El diseño experimental constituye uno de los pilares metodológicos más sólidos dentro de la *investigación científica*, ya que permite establecer relaciones causales mediante la manipulación de variables independientes y la observación sistemática, a través de la medición de las variables en las unidades de estudio (efectos), con un adecuado control de las variables intervinientes.

Desde una perspectiva amplia, los experimentos han acompañado el desarrollo del conocimiento científico desde sus primeras formulaciones en la antigüedad; sin embargo, adquirieron un carácter formal y riguroso a partir del surgimiento del método científico moderno, especialmente durante los siglos XVII y XVIII. Con el avance de la Estadística en el siglo XX, figuras como Ronald A. Fisher fortalecieron el diseño experimental al incorporar principios básicos que resultan fundamentales en la actualidad, como la aleatorización y replicación de tratamientos; así como, el control local del error experimental. Estos aportes sentaron las bases de los modelos experimentales contemporáneos y definieron la estructura lógica que guía la investigación empírica en múltiples áreas de conocimientos.

Dentro de esta evolución histórica, las características esenciales del diseño experimental se consolidaron como elementos indispensables para garantizar la validez de los resultados. Entre ellas, destacan el control de las condiciones, la manipulación de una o varias variables independientes, la observación precisa de los cambios en la variable dependiente y la posibilidad de repetir el procedimiento para obtener

conclusiones consistentes. A su vez, estas características se articulan con etapas claramente definidas, que incluyen la formulación del problema, la identificación de variables, la elección del tipo de experimento, la selección de la muestra, la ejecución del procedimiento estadístico y el análisis e interpretación de los resultados obtenidos. Cada una de estas fases contribuye a la transparencia metodológica, garantizar reproducibilidad y asegurar que los hallazgos puedan ser interpretados y evaluados críticamente.

El propósito de esta obra es servir como texto de consulta indispensable para estudiantes de las carreras de agronomía, medicina veterinaria, acuicultura y agropecuaria pertenecientes a la Facultad de Ciencias Agropecuarias, Universidad Técnica de Machala (UTMACH), provincia El Oro, Ecuador, ya que no se cuenta con un libro de texto que abarque todos los contenidos mínimos planificados en el Syllabus de la asignatura de *diseño experimental*, relacionados con los experimentos unifactoriales y factoriales.

En el libro se combinan, de manera simultánea, los elementos teóricos (se detallan los fundamentales) y prácticos (se profundiza en el desarrollo de ejercicios prácticos con el manejo de software estadístico) que permiten a través de la esquematización de los procedimientos lograr un mejor entendimiento por el público objetivo, además, se detallan elementos relacionados con la conducción del experimento, recolección de datos, creación de las matrices de datos y procedimientos estadísticos relacionados con estudios experimentales.

La importancia del diseño experimental radica en su capacidad para generar conocimiento objetivo; así como, su flexibilidad para adaptarse a diversas disciplinas y problemáticas. Comprender sus fundamentos históricos, lógicos y metodológicos permite al investigador seleccionar y aplicar procedimientos experimentales acordes con sus propósitos, fortaleciendo la calidad y la credibilidad de sus resultados. Así, el diseño experimental se presenta no solo como una técnica, sino como una estrategia metodológica integral que articula teoría, práctica y evidencia verificable para avanzar en la construcción del conocimiento científico.

La finalidad del libro es que los estudiantes y profesionales de las ciencias agropecuarias sean capaces de reconocer los elementos teóricos y prácticos que posibiliten el diseño y conducción de experimentos donde el investigador manipula de forma deliberada uno o más factores de estudio, así como, el desarrollo de procedimientos estadísticos univariados, bivariados y multivariados a partir de datos recolectados en investigaciones desarrolladas en las ciencias agropecuarias.

El objetivo de la obra es Diseñar y conducir experimentos con uno o más factores de estudio con intervención deliberada del investigador en las unidades de estudio, mediante la aplicación de los principios básicos de una buena experimentación y su articulación metodológica y Estadística, que posibilite la verificación de hipótesis científicas planteadas en las ciencias agropecuarias.

El alcance del libro es cubrir las unidades 1 y 2 del syllabus de la asignatura de diseño experimental en la Facultad de Ciencias Agropecuarias, Universidad Técnica de Machala.

Los estudiantes tendrán la competencia de planificar y ejecutar experimentos con uno o más factores de estudio a partir del cumplimiento de los principios y requisitos básicos de la experimentación agropecuaria, además de interpretar y comunicar los resultados obtenidos con un fundamento científico, que posibiliten, mediante trabajo en equipo, el desarrollo de proyectos de investigación encaminados a solucionar problemas en el entorno productivo y generar innovación.

La competencias a lograr es potenciar en los estudiantes el desarrollo de habilidades duras y blandas, competencias clave establecidas en el actual modelo educativo de la UTMACH, que fortalecen el perfil de egreso de las carreras ofertadas en la Facultad de Ciencias Agropecuarias (agronomía, medicina veterinaria, acuicultura y agropecuaria) de la UTMACH y posibiliten la articulación de la metodología de la investigación, el diseño del estudio y la Estadística en investigación científica, la escritura científica y la interpretación de fenómenos biológicos con un fundamento científico y estadístico.

ÍNDICE

PREFACIO ..... **i**

ÍNDICE ..... **v**

ÍNDICE DE TABLAS ..... **ix**

ÍNDICE DE FIGURAS ..... **xii**

CAPÍTULO I **17**

1 EL DISEÑO EXPERIMENTAL EN INVESTIGACIÓN CIENTÍFICA **17**

1.1 Tipos de investigación 20

1.1.1 Según su propósito 20

1.1.2 Según el enfoque 21

1.1.3 Según la fuente de los datos 22

1.1.4 Según los niveles de la investigación 22

1.2 Tipos de estudio 24

1.2.1 Según la intervención del investigador 24

1.2.2 Según la planificación de las mediciones 24

1.2.3 Según el número de mediciones 25

1.2.4 Según el número de variables procesadas estadísticamente 25

1.3 El diseño de experimentos en la investigación agrícola 26

1.4 Generalidades del diseño experimental 29

1.4.1 Contenidos 29

1.4.2 Objetivos de aprendizaje 29

1.4.3 Introducción 30
1.4.4 Breve historia del diseño experimental 31
1.4.5 Características de los diseños experimentales unifactoriales 32
1.4.6 Clasificación de los tipos de experimentos de acuerdo a diferentes criterios 34
1.4.7 Etapas o fases para el establecimiento de un diseño experimental 36
1.4.8 Conceptos básicos en diseño experimental 39
1.4.9 Ejemplo de un experimento con testigo absoluto y con testigo relativo 45
1.4.10 Orígenes del error experimental 49
1.5 Práctica 1. Introducción al manejo de Software estadístico Jamovi 50
1.5.1 Contenidos 50
1.5.2 Objetivos de aprendizaje 50
1.5.3 Introducción 50
1.5.4 Desarrollo 51
1.5.5 Ventajas 52
1.5.6 Procedimiento para iniciar la ejecución del programa Jamovi. . 52
1.5.7 Cuadros de diálogos iniciales 54
1.5.8 Otras opciones de Jamovi 79
1.5.9 Ejercicio 1. Importar datos de Excel 81
1.6 Principios y requisitos básicos del diseño experimental 84

1.6.1 Contenidos 84
1.6.2 Objetivos de aprendizaje 84
1.6.3 Introducción 84
1.7 Práctica 2. Edición y transformación de variables con Jamovi. 100
1.7.1 Contenidos 100
1.7.2 Objetivos de aprendizaje 100
1.7.3 Introducción 100
1.7.4 Desarrollo 101
1.7.5 Ejercicio 3. Calcular una nueva variable. 103
1.7.6 Ejercicio 4. Transformación de variables 108
CAPÍTULO II 116
2 APLICACIONES PRÁCTICAS DE LA ESTADÍSTICA 116
2.1 Estadística descriptiva 117
2.1.1 Contenidos 117
2.1.2 Objetivos de aprendizaje 118
2.1.3 Introducción 118
2.1.4 Conceptos fundamentales 123
2.1.5 Otras opciones para la construcción de gráficos univariados con Jamovi 161
2.2 Práctica 3. Descripción tabular y representación gráfica de variables numéricas y categóricas 164
2.2.1 Ejercicio 3.1 Creación de matrices de datos con Jamovi. 165

2.2.2 Ejercicio 5 Estadísticos descriptivos en variables categóricas. 170

2.3 Estadística inferencial 175

2.3.1 Contenidos 175

2.3.2 Objetivos 176

2.3.3 Introducción 176

2.4 Práctica 4. Pruebas paramétricas en investigación experimental. 231

2.4.1 Contenidos 231

2.4.2 Objetivo de aprendizaje..... 231

2.4.3 Introducción 231

2.4.4 Ejercicio 6. Prueba t de Student para tratamientos o grupos independientes. 232

2.4.5 Ejercicio 7. ANOVA de un factor intertratamientos o grupos.. 243

ANEXOS 254

BIBLIOGRAFÍAS 259

ÍNDICE DE TABLAS

Tabla 1. Ejemplos. 44

Tabla 2. Importar datos de Excel. 82

Tabla 3. Replicación o repetición de tratamientos. 89

Tabla 4.Resumen del contexto. 93

Tabla 5.Tratamiento. 96

Tabla 6.Calcular una nueva variable. 104

Tabla 7.Transformación de variables. 109

Tabla 8.Requerimiento del ejercicio. 109

Tabla 9. Técnicas de recolección de datos en función del tipo de dato obtenido en el estudio 124

Tabla 10. Transformación de materia prima en productos. 125

Tabla 11. Rangos del CV que establecen la calidad o precisión de un experimento en ciencias agropecuarias. 147

Tabla 12. Resumen de los principales gráficos univariados que se realizan por tipo de variable. 155

Tabla 13. Datos recolectados. 167

Tabla 14. Tabla de estadísticos descriptivos de la variable peso de caninos 169

Tabla 15. Estadísticos descriptivos en variables categóricas. 171

Tabla 16. Distribución de frecuencias de la variable peso de caninos categorizada. 174

Tabla 17. Formas de escribir las hipótesis con lenguaje estadístico cuando se dispone de dos o más tratamientos o grupos. 186

Tabla 18. Situaciones posibles asociadas a las distintas decisiones del investigador al desarrollar un procedimiento de prueba de hipótesis. 190

Tabla 19. Principales pruebas paramétricas y no paramétricas. 192

Tabla 20.Calcular las medias muestrales y las varianzas de cada finca. 211

Tabla 21.Resultados de las puntuaciones. 218

Tabla 22. Construcción de la tabla de ANOVA de un factor intertratamientos o grupos. 224

Tabla 23.Pruebas post hoc implementadas en Jamovi. 227

Tabla 24. Niveles de clasificación más empleados para el tamaño del efecto. 229

Tabla 25.Prueba t de Student para tratamientos o grupos independientes. 233

Tabla 26. Prueba de Shapiro-Wilk. 236

Tabla 27.Prueba de Levene. 237

Tabla 28.Prueba t de Student para tratamientos o grupos independientes. 239

Tabla 29.Estadísticos descriptivos de cada tratamiento. 239

Tabla 30. Datos recolectados. 243

Tabla 31.Prueba de Shapiro-Wilk. 246

Tabla 32.Prueba de Levene. 248

Tabla 33.ANOVA de un factor intertratamientos o grupos (Fisher). . 249

Tabla 34. Estadísticos descriptivos de cada tratamiento. 250

Tabla 35. Prueba post hoc (Tukey) para fórmulas de fertilizantes y rendimiento agrícola (t ha-1). 251

ÍNDICE DE FIGURAS

Figura 1. Componentes básicos de la investigación científica basada en evidencia. 19
Figura 2. Niveles de la investigación científica considerados en la evolución de una línea de investigación..... 20
Figura 3.Factores que intervienen en un experimento verdadero..... 41
Figura 4. Cuadro de diálogo en azul. 53
Figura 5. Menú inicio. 53
Figura 6. Archivos abiertos recientemente. 54
Figura 7. La pantalla de Jamovi-Untitled. 54
Figura 8.Cuadro de dialogo. 55
Figura 9.Utilizada para abrir archivos que se encuentran dentro de tu ordenador. 56
Figura 10.Se utiliza para exportar el archivo. 56
Figura 11.Extensión nativa del software. omv. 57
Figura 12.Pestaña de variables. 58
Figura 13.Cuadro de diálogo. 58
Figura 14.Variable numérica. 60
Figura 15.Eliminar la variable. 60
Figura 16.Subconjuntos de casos. 61
Figura 17.Cuadro de dialogo. 62
Figura 18.Pestaña análisis. 64
Figura 19.Gráfico de encuestas. 65
Figura 20.Pruebas t de Student. 65
Figura 21.Regresión en Jamovi. 66
Figura 22.Cuadro de diálogo. 68
Figura 23.Modelo estadístico. 70

Figura 24.Análisis de Componentes Principales. 71
Figura 25.Estadística de variables. 72
Figura 26.Módulo JJStatsPlot. 74
Figura 27.JJStatsPlot. 75
Figura 28.Software estadístico R. 76
Figura 29.SnowCluster módulo que se activa de forma predeterminada. 77
Figura 30.Barra de menú. 79
Figura 31.Minimizar. 80
Figura 32.Guardar los cambios. 80
Figura 33.Filtros. 81
Figura 34. Software estadístico moderno. 83
Figura 35.Tabla de números aleatorios. 95
Figura 36.Unidades experimentales. 96
Figura 37.Area experimental. 97
Figura 38.Esquema de aleatorización. 98
Figura 39.Características de la variable. 102
Figura 40.Variable de Jamovi. 103
Figura 41.Menú variables. 104
Figura 42.Menú datos. 105
Figura 43.Variables originales. 106
Figura 44.Transformaciones matemáticas o cálculos. 106
Figura 45.Transformación de variables. 108
Figura 46.Recodificaciones. 108
Figura 47.Matriz de datos 110
Figura 48.Variable de origen. 111
Figura 49.Categoría de IMC 111

Figura 50.Nueva Transformación. 112
Figura 51.Transformación. 112
Figura 52.Condición de recodificación. 113
Figura 53.Categorías de IMC. 113
Figura 54.Cuadro de diálogo. 114
Figura 55.Función matemática. 114
Figura 56.Rangos y categorías..... 115
Figura 57. Etapas del proceso de investigación científica donde la Estadística se encuentra presente. 123
Figura 58.Criterios de clasificación de las variables y donde se utilizan. 129
Figura 59. Atributos de medición de las variables categóricas y numéricas. 133
Figura 60. Esquema del proceso experimental con la clasificación de las variables por su relación causa-efecto..... 134
Figura 61.Media aritmética. 140
Figura 62. Distribución de los cuartiles en función de los porcentajes de valores que involucran. 150
Figura 63. Tipos de asimetría. 152
Figura 64. La curtosis. 153
Figura 65.Análisis..... 156
Figura 66.A. Histograma simple de la variable peso de estudiantes. B. Histograma con la curva de distribución de los datos. 158
Figura 67.Diagrama de caja y sesgos 1-D de la variable peso de estudiantes. 159
Figura 68. Gráfico de barras simples de la variable razas de ganado vacuno. 160

Figura 69. Gráfico circular de la variable IMC. 161
Figura 70. Módulo JJStatsPlot. 162
Figura 71.Estadísticos descriptivos en variables numéricas. 166
Figura 72. Software estadístico. 168
Figura 73. Asimetría. 170
Figura 74.Gráfico circular. 173
Figura 75.Frecuencia relativa. 173
Figura 76. Representación gráfica de los resultados obtenidos en la variable peso de caninos (categorizada). 174
Figura 77. Flujograma que explica el camino a seguir en función de la intención analítica del investigador y el procedimiento estadístico. 178
Figura 78. Flujograma que explica el camino a seguir en función de la intención analítica del investigador y el procedimiento estadístico. 182
Figura 79.Distribución muestral del estadístico z; prueba de una cola a la izquierda; nivel de significancia de 0,05. Fuente: Lind et al. (2012). 184
Figura 80.Distribución muestral del estadístico z; prueba de una cola a la derecha; nivel de significancia de 0,05. Fuente: Lind et al. (2012). 184
Figura 81. Regiones de aceptación y rechazo de una prueba de dos colas con un nivel de significancia de 0,05. Fuente: Lind et al. (2012). 185
Figura 82. Representación gráfica del nivel de significación (α) en pruebas de hipótesis, indicando el umbral de error tipo I definido previamente por el investigador (α = 0,05). 195

Figura 83. Esquema de la tabla del análisis de varianza de un factor intertratamientos o grupos. 197
Figura 84.Representación gráfica de las propiedades de la distribución normal. 206
Figura 85. Comportamiento de la variabilidad de los valores en una distribución normal. 207
Figura 86.Verificación gráfica de la homogeneidad de varianzas entre dos tratamientos o grupos independientes. A. Varianzas diferentes. B. Varianzas iguales. 209
Figura 87.Gráfico cartesiano que representa las fuentes de variabilidad en un ANOVA de un factor intertratamientos o grupos. 220
Figura 88. Pruebas t. 235
Figura 89.Efecto de preparados antivirales en la superficie de lesiones locales en hojas de tabaco. 242
Figura 90.Procedimiento estadístico..... 245
Figura 91.Efecto de la fórmula de fertilizante en el rendimiento agrícola del cultivo de tomate. 252
Figura 92. Valores críticos de la Distribución t de Student. 255
Figura 93.Valores críticos de la distribución F para los niveles de significación más frecuentes. 258

CAPÍTULO I

1 EL DISEÑO EXPERIMENTAL EN INVESTIGACIÓN CIENTÍFICA

La experimentación en ciencias agropecuarias se sustenta fundamentalmente en los principios formulados por Fisher y desarrollados posteriormente por Snedecor, Cochran y Gomez, quienes consolidaron el uso del diseño experimental en ensayos agrícolas.

En la Figura 1.1 se representa un modelo conceptual integrador donde la línea de investigación (LIN) constituye el marco general dentro del cual se desarrolla el proceso científico apoyado por cuatro componentes fundamentales: metodología de la investigación (responde a la línea de investigación), diseño del estudio (coherente con la LIN), estadística (genera evidencia para responder a sus preguntas científicas y a los objeticos planteados) e inteligencia artificial (IA), la cual se utiliza como herramienta de apoyo dentro de ese contexto. En el centro de la interacción de estos elementos se encuentra la investigación científica (ICE) que funciona como el núcleo o eje articulador del proceso investigativo y donde la metodología define el camino científico, el diseño operacionaliza la investigación, la estadística valida e interpreta los resultados y la IA potencia la generación y el análisis del conocimiento.

La metodológica de la investigación (ciencia del método científico que posibilita establecer la estrategia metodológica) constituye el conjunto de enfoques, principios, métodos y procedimientos que se utilizan para

generar conocimiento científico. Representa el componente que estructura el proceso investigativo y garantiza su rigor científico, ya que orienta la formulación del problema, define la hipótesis y los objetivos, determina los métodos a utilizar para la recolección de datos y establece la lógica de la validación científica.

El diseño del estudio (pueden ser experimental o no experimental) define como se obtendrán los datos de las variables de interés y como se controlan los posibles sesgos que se puedan presentar en la investigación).

La Estadística, que puede ser descriptiva (permite organizar matrices de datos, calcular medidas de resumen de datos y caracterizar variables numéricas o categóricas) e inferencial (contrasta hipótesis mediante el procedimiento de prueba de hipótesis, estima parámetros, evalúa relaciones probabilísticas y causales entre variables y generaliza los resultados obtenidos en la muestra a la población objeto de estudio (POE). Su importancia radica en que funciona como mecanismo utilizado para transformar datos en evidencia científica, o sea, es el puente entre los datos obtenidos en el estudio y las conclusiones científicas derivadas de la investigación. Es una herramienta que apoya la toma de decisiones en un entorno de incertidumbre y y facilita la interpretación de los resultados alcanzados en el fenómeno estudiado con un argumento matemático.

La IA funciona como una herramienta transversal y complementaria de apoyo a la investigación ya que no reemplaza el método científico y tampoco la capacidad analítica del investigador. La IA puede apoyar en

la búsqueda bibliográfica, revisión documental, procesamiento de los datos, generación de ideas, identificación de patrones complejos, modelado predictivo y automatización de tareas repetitivas, por ello, contribuye a aprender, investigar y crear.

Los estudiantes-investigadores deben dominar estos componentes y desarrollarlos de forma articulada, lo cual les permitirá formar competencias investigativas y fortalecer la LIN, además de prepararlos para el desarrollo del proceso de titulación (asignaturas de titulación I y titulación II) y culminar con éxito sus carreras universitarias.



Figura 1. Componentes básicos de la investigación científica basada en evidencia.

La LIN funciona como el marco estratégico que articula y proporciona coherencia al proceso científico. En la UTMACH las LIN se encuentran construidas y aprobadas por cada dominio académico, por ello, cada investigador debe fortalecerlas, y así, impulsar la producción o servicios en la zona de influencia de la Institución de Educación Superior. Desde el punto de vista estadístico cada LIN transita por diferentes niveles de

la investigación (NIV) (fases, momentos o etapas) en los cuales el procesamiento estadístico de los datos se realiza con diferentes pruebas univariadas, bivariadas y multivariadas (Figura 2).



Figura 2. Niveles de la investigación científica considerados en la evolución de una línea de investigación.

A manera de resumen a continuación se clasifican los tipos de investigación científica

1.1 Tipos de investigación

1.1.1 Según su propósito

Pueden ser:

- **Básica:** su finalidad es cognoscitiva, o sea, generar conocimiento sin una aplicación inmediata. **Ejemplo:** determinar los mecanismos fisiológicos y moleculares que explican la tolerancia a la sequía en diferentes variedades de garbanzo bajo condiciones

controladas de laboratorio. Se encuentra en los niveles de la investigación exploratorio, descriptivo, relacional, explicativo y predictivo.

- **Aplicada:** el investigador interviene para resolver un problema práctico y transformar de forma positiva la realidad de las personas y su entorno. **Ejemplo:** evaluar el efecto de diferentes dosis de compost elaborado a partir de residuos de banano en el rendimiento agrícola del cultivo de garbanzo en la Granja Experimental Santa Inés. Se encuentra en el nivel de la investigación aplicativo.

1.1.2 Según el enfoque

Puede ser:

- **Cualitativa:** se desarrollan teorías emergentes, descripciones profundas, narrativas. Busca comprender experiencias, percepciones o significados. **Ejemplo:** Análisis mediante entrevistas en profundidad de las percepciones de los agricultores de la provincia de El Oro sobre las barreras que se presentan para la adopción de prácticas de agricultura sostenible. Se encuentra en el nivel de la investigación exploratorio.
- **Cuantitativa:** resultados numéricos, gráficos, tablas y significancia estadística. Se efectúa mediciones de las variables y se realiza procesamiento estadístico. **Ejemplo:** Determinar la relación probabilística entre la dosis de fertilizante orgánico aplicada y el rendimiento agrícola del cultivo de maíz mediante pruebas de significación estadística y análisis de regresión lineal

simple. Se encuentra en los niveles de la investigación descriptivo, relacional, explicativo, predictivo y aplicativo.

1.1.3 Según la fuente de los datos

Pueden ser:

- **Documental:** información disponible en documentos, bases de datos, artículos científicos, archivos históricos, historias clínicas en una clínica veterinaria. **Ejemplo:** revisión sistemática de los estudios publicados en revistas de alto impacto entre 2015 y 2025 sobre índices de calidad del suelo en agroecosistemas tropicales. Se encuentra en el nivel de la investigación exploratorio.
- **De campo:** demanda de la observación directa de los objetos o sujetos y recolección de datos en el lugar de estudio. **Ejemplo:** medición de la altura de plantas y el rendimiento agrícola del cultivo de soya en parcelas experimentales sometidas a diferentes distancias de siembra. Se encuentra en los niveles de la investigación descriptivo, relacional, explicativo, predictivo y aplicativo.

1.1.4 Según los niveles de la investigación

Pueden ser:

- **Exploratorio:** se analizan fenómenos poco estudiados mediante la revisión de información publicada. **Ejemplo:** explorar la utilización de biofertilizantes elaborados por pequeños agricultores de El Oro a partir de residuos vegetales y sus posibles beneficios para la producción agrícola.

- **Descriptivo**: describe fenómenos en una circunstancia temporal y espacial determinada. Caracteriza una población, situación o fenómeno. **Ejemplo:** describir las propiedades físicas, químicas y biológicas de suelos cultivados con banano en distintos agroecosistemas de la provincia de El Oro.
- **Relacional**: cuantifica la relación entre dos variables medidas en diferentes escalas. Estudia relaciones probabilísticas. Determina el grado de asociación o correlación o dos variables. **Ejemplos:** Establecer la correlación el contenido de materia orgánica del suelo y el rendimiento agrícola del cultivo de maíz. **Ejemplo:** Establecer la asociación entre la presencia o ausencia de *Babesia canis* y el tipo de raza de caninos atendidos en la Clínica de Especialidades Veterinarias de la UTMACH.
- **Explicativo:** estudios de causalidad donde se explica el comportamiento de una o más variables respuestas en función de otra u otras variables causas manipuladas. Su finalidad es detectar relaciones de causa y efecto. **Ejemplo:** determinar el efecto causal de diferentes dosis de fertilización nitrogenada en el crecimiento y rendimiento agrícola del cultivo de ajonjolí.
- **Predictivo:** se pronóstica la ocurrencia de eventos futuros mediante la construcción de modelos predictivos. **Ejemplo:** desarrollar un modelo predictivo del rendimiento agrícola del fréjol utilizando variables climáticas, propiedades edáficas e índices de vegetación obtenidos por drones.
- **Aplicativo:** se mejoran las condiciones de vida de las personas y su entorno. Enmarca la generación de innovación técnica, tecnológica, artesanal o científica. **Ejemplo:** Diseñar e

implementar un sistema de manejo integrado de fertilización orgánica que posibilite incrementar la productividad agrícola y sostenibilidad de los agroecosistemas.

A manera de resumen a continuación se clasifican los tipos de estudio en investigación científica

1.2 Tipos de estudio

1.2.1 Según la intervención del investigador

Pueden ser:

- **No experimental**: el investigador se limita a observar la evolución natural de los eventos. No se presenta manipulación deliberada de variables. **Ejemplo:** Analizar la relación entre el contenido de carbono orgánico del suelo y la productividad de fincas bananeras utilizando datos observados en condiciones naturales.
- **Experimental:** intervención deliberada por parte del investigador quien manipula una o más variables. **Ejemplo:** Comparar el efecto de cuatro dosis de compost sobre el rendimiento agrícola del cultivo de alfalfa mediante un diseño en bloques completamente al azar.

1.2.2 Según la planificación de las mediciones

Pueden ser:

- **Prospectivo:** se realizan mediciones planeadas por parte del investigador antes de iniciar el estudio. **Ejemplo:** Diseñar un experimento unifactorial para evaluar durante un ciclo agrícola

completo el efecto de diferentes distancias de siembra sobre el crecimiento vegetativo del cultivo de maíz.

- **Retrospectivo:** con mediciones no planeadas ni realizadas por el investigador, quien utiliza datos ya existentes. No controla los sesgos de medición. **Ejemplo:** Analizar registros de rendimiento agrícola del cultivo de plátano recopilados por una empresa exportadora entre 2018 y 2025 con la finalidad de identificar tendencias productivas.

1.2.3 Según el número de mediciones

Pueden ser:

- **Transversal:** se efectúa una sola medición por cada unidad de estudio o de muestreo. **Ejemplo:** evaluar en una única fecha la calidad física, química y biológica de los suelos de 100 fincas agropecuarias integrales.
- **Longitudinal:** se efectúan múltiples mediciones en las unidades de estudio o de muestreo en un periodo de tiempo determinado. **Ejemplo:** monitorear trimestralmente durante tres años consecutivos los cambios ocurridos en los indicadores de calidad del suelo bajo diferentes sistemas de manejo agrícola.

1.2.4 Según el número de variables procesadas estadísticamente

Pueden ser:

- **Descriptivo:** se analizan las variables de interés de una en una. **Ejemplo:** Determinar el promedio, desviación típica o estándar de

los datos y distribución del pH en los suelos agrícolas de una determinada zona.

- **Analítico:** analiza dos o más variables analíticas o de interés de forma simultánea. **Ejemplo:** Evaluar la influencia conjunta de la materia orgánica, la densidad de la fase sólida y el agua aprovechable del suelo en el rendimiento agrícola del cultivo de garbanzo mediante regresión lineal múltiple.

1.3 El diseño de experimentos en la investigación agrícola

La experimentación pertenece al método científico empírico-analítico (es fáctico, ya que se basa en hechos comprobables; verificable, debido a que se utilizan instrumentos de medición; progresivo y autocorrectivo, por actualizarse de manera paulatina y constante, y depende del muestreo, ya que pueden obtenerse resultados no confiables por una recolección defectuosa de datos) y el más utilizado en las ciencias naturales, emplea la observación como técnica de recolección de datos para responder a las hipótesis planteadas y poder generar nuevos conocimientos (Concepto, 2025).

En la experimentación el investigador se debe enforcar en lograr desarrollar un proceso sencillo que facilite aplicar el diseño adecuado y efectuar una conducción de forma detallada y minuciosa.

En el diseño de una investigación (sea experimental u observacional) se desglosan las estrategias básicas que el investigador adopta para generar información exacta e interpretable (los datos recolectados con instrumentos de medición precisos y exactos se organizan en una matriz de datos, se procesan estadísticamente y se convierten en información),

y constituye la estrategia general que se establece para responder a la hipótesis planteada; que a la postre conducen a desarrollar nuevas teorías o tecnologías que impulsan el desarrollo o la productividad de la producción o los servicios.

El diseño de experimentos se define como un proceso mediante el cual se planifican los pasos o etapas necesarios que se deben cumplir; así como, la ordenación que deben alcanzar para la recolección de datos, procesamiento estadístico y posterior análisis de la información que se requiere estudiar en un problema de investigación. Su objetivo es obtener el máximo de datos requeridos con el mínimo costo y la máxima eficiencia. Se deben realizar mediciones con precisión y exactitud; y proporcionar la mayor información con un número menor de datos. Por todo ello, es necesario tener seguridad de lo que se está buscando (Díaz, 2009).

El diseño experimental y el procesamiento de datos en una investigación se encuentran intrínsecamente interconectados. Las razones por las cuales no pueden ir separados se enumeran a continuación:

1. **Definición de objetivos y metodología**: El diseño de una investigación define los objetivos y la metodología a utilizar. Sin un diseño claro, el análisis de los resultados no tendría una base sólida y confiable.

2. **Recolección de datos**: El diseño de la investigación determina cómo se recolectarán los datos. Si el diseño es defectuoso, la calidad y relevancia de los datos recolectados estarán

comprometidas, lo cual afecta el análisis posterior de los resultados.

3. **Validez y fiabilidad**: Un buen diseño asegura la validez y fiabilidad de los resultados. El procesamiento de datos sin un diseño sólido puede conducir a interpretaciones erróneas o sesgadas.

4. **Control de variables**: El diseño de investigación incluye la identificación y control de variables. Sin un diseño adecuado, el procesamiento de datos puede afectarse por variables no controladas que introducen errores en los resultados.

5. **Interpretación de resultados**: Un diseño bien elaborado permite una interpretación adecuada de los resultados obtenidos en el procesamiento de los datos realizado. Sin esta estructura, los resultados pueden ser malinterpretados o carecer de significado.

6. **Reproducibilidad**: Un diseño bien documentado y ejecutado asegura que otros investigadores puedan reproducir el experimento y validar los resultados obtenidos en el procesamiento de datos.

7. **Repetibilidad**: los resultados obtenidos en un experimento; deben ser similares a los alcanzados en experimentos realizados, por otros investigadores, en las mismas condiciones.

El diseño experimental y el procesamiento estadístico de datos son procesos que se articulan uno con el otro. Un diseño robusto proporciona la estructura necesaria para un procesamiento estadístico de datos

preciso y significativo. Los procedimientos estadísticos permiten confirmar o validar los pasos desarrollados en el experimento y la relación entre las variables.

En el desarrollo de diseños experimentales se debe tener presente lo siguiente:

1. Mantener siempre la máxima sencillez (FAO, 2010).
2. Garantizar confiabilidad en los datos. Pocos datos válidos, son más útiles que una gran cantidad de datos que no se puede utilizar porque no son confiables.
3. Las mediciones a realizar en campo deben ser lo más sencillas posibles y orientarse a objetivos limitados y alcanzables.
4. Considerar la medición de las variables que responden a los objetivos planificados.
5. Considerar los recursos disponibles, teniendo en cuenta los gastos iniciales y de funcionamiento, además, del personal requerido para la conducción del experimento.

1.4 Generalidades del diseño experimental

1.4.1 Contenidos

Historia, características, ventajas y limitaciones del DEXPE. Tipos de experimentos. Fases para establecer un experimento. Conceptos básicos en diseño experimental.

1.4.2 Objetivos de aprendizaje

1. Reconocer las principales generalidades del diseño experimental utilizado en investigación científica.

2. Identificar las etapas que abarca el establecimiento de un experimento en investigación científica.
3. Definir los principales conceptos básicos y la terminología técnica adecuada en el diseño de experimentos.

1.4.3 Introducción

Al profundizar en los conceptos básicos del DEXPE, se encuentran términos como unidad experimental, unidad de estudio, factores de estudio, factores controlados, factores no controlados, variable independiente, dependiente y de control, validez interna y externa, aleatorización, replicación, tratamiento experimental y tratamiento control, todos los cuales conforman el lenguaje técnico necesario para comprender y llevar a cabo investigaciones experimentales de manera adecuada. La claridad conceptual y operativa respecto a estos elementos proporciona al investigador las herramientas para diseñar estudios más sólidos y responder a preguntas científicas con mayor precisión.

El DEXPE ofrece ventajas notables, dentro de las cuales se encuentran la capacidad de establecer causalidad, el control riguroso de factores externos y la posibilidad de replicación; sin embargo, presenta limitaciones, principalmente relacionadas con la dificultad de reproducir condiciones estrictamente controladas en contextos naturales, los costos asociados, o la inconveniencia ética de manipular ciertas variables. Estas consideraciones han dado lugar a una diversidad de enfoques y tipos de experimentos, entre los cuales se encuentran los diseños puros o verdaderos, los cuasiexperimentales y los preexperimentales. Cada tipo responde a diferentes niveles de control y rigor, lo que permite adaptar

el método experimental a una amplia variedad de escenarios investigativos.

1.4.4 Breve historia del diseño experimental

El diseño estadístico de experimentos (introducido por primera vez en las ciencias agrícolas por Ronald Fisher en la primera mitad del siglo XX (1920-1930) en Inglaterra, donde fue el responsable de las estadísticas y el procesamiento de datos en la estación Agrícola Experimental de Rothamsted en Londres) fue ut20ilizado para conseguir un aprendizaje acelerado.

El libro de Fisher (*The design of experiments*) (Fisher, 1935) influyó decisivamente en la investigación agrícola ya que:

- Aportó métodos que permiten evaluar los resultados de experimentos con muestras pequeñas.
- Entendió que la forma de realizar el experimento puede obstaculizar el procesamiento estadístico de los datos y el análisis e interpretación de los resultados experimentales.
- Incorporó los principios estadísticos en el diseño de investigaciones experimentales (análisis de varianza).
- Desarrolló métodos para desarrollar experimentos destinados a investigar la influencia simultanea de varios factores.

Fisher, en su interacción con otros investigadores, establecieron los tres principios básicos del diseño de experimentos, la aleatorización de tratamientos, replicación de tratamientos y el control local de error

experimental (incluye utilización de técnica de bloqueo si fuese necesario).

El diseño de experimentos entre 1950 y 1980 se convirtió en una herramienta de aplicación frecuente en las áreas de investigación y desarrollo; y en la actualidad se utiliza en todas las áreas del conocimiento.

1.4.5 Características de los diseños experimentales unifactoriales

- **Equivalencia estadística:** garantizar la homogeneidad de los objetos (plantas, animales, agua, suelo, etc.) o sujetos (personas: productores, representantes de mascotas, trabajadores, etc.) que se ubican en los diversos grupos formados al azar (evita que las desigualdades en los resultados no sean provocadas por diferencias iniciales en los grupos).
- **Comparación de dos o más tratamientos**: siempre que se realiza un experimento verdadero la finalidad estadística es comparar los tratamientos en función las variables medidas en las unidades de estudio o de muestreo).
- **Manipulación directa de una variable independiente:** es el factor de estudio que se manipula en diferentes niveles (valores) o versiones (atributos) que el investigador asigna (experimento verdadero) de forma aleatoria.
- **Medición de cada variable dependiente:** el resultado de la investigación debe ser medido; constituye el efecto que se mide en las unidades de estudio o de muestreo y que responde a los objetivos planificados en el estudio. En cada experimento pueden ser medidas una o más de una variable.

- **Uso de Estadística inferencial:** diseño que permita generalizar los resultados a la población objeto de estudio a partir de los datos obtenidos en la muestra, la cual debe ser representativa.

Ventajas del DEXPE

- La manipulación del factor de estudio y el control de las variables que intervienen en el experimento clarifican la dirección y naturaleza de la causa.
- La aleatorización de los tratamientos minimiza la acción de las variables intervinientes (extrañas o de ruido), sean confusoras o fatores no controlados.
- Proporcionan una respuesta más rápida y barata a una hipótesis que un estudio observacional.
- Garantizan evidencia sólida en base a inferencias causales.

Limitaciones del DEXPE

- **De tipo ambiental:** son propias del ambiente y pueden afectar el resultado de la investigación; por ello, los resultados no pueden generalizarse a otras investigaciones.
- **En la medida:** a veces en estudios con personas es difícil llegar a conocer la realidad, por las dificultades que se presentan en la medición, debiéndose utilizar instrumentos de medición, mecánicos o documentales, válidos y confiables.
- **De tipo moral:** cuando se realizan investigaciones con seres humanos las condiciones de tipo moral afectan a los sujetos que participan.

- **Derivadas del objeto o sujeto de estudio** (todas las mediciones deben ser verificable).

Elementos clave del Diseño Experimental

- **Factores y tratamientos:** un factor de estudio que se manipula de forma deliberada por el investigador y los niveles o versiones son los distintos tratamientos o condiciones de ese factor de estudio.
- **Mejorar la precisión:** asegurar que las diferencias observadas en las variables respuestas se deban al efecto de los tratamientos aplicados y no a otras fuentes de variación no controladas.
- **Maximizar la eficiencia:** obtener la mayor cantidad de datos con el menor costo y esfuerzo posible.
- **Validez del estudio:** un buen diseño experimental permite realizar conclusiones válidas y generalizables sobre la relación causa-efecto entre las variables estudiadas. Este es un fundamento esencial para la investigación científica y es aplicable en diversas disciplinas, incluidas las ciencias agropecuarias.

1.4.6 Clasificación de los tipos de experimentos de acuerdo a diferentes criterios

1. Según el grado de control de las variables por parte del investigador

- **Preexperimento.** Carece de grupo control; sirven como primer acercamiento a ciertos tópicos, es decir, en investigaciones exploratorias y descriptivas. En este caso, se tiene poco control de las variables y no se puede estar seguro de que el resultado

obtenido se deba única y exclusivamente a una de ellas (se utilizan en pruebas piloto).

- **Cuasiexperimento.** Aquellos en que se tienen dos o más grupos de comparación; sin embargo, la constitución de los mismos es previa al experimento, se caracterizan especialmente por la ausencia de asignación aleatoria (Manterola y Otzen, 2015), son ordenados a priori para demostrar algo, provienen del ámbito educativo y de la psicología (en la actualidad se denominan experimentos naturales), sin embargo, pueden aplicarse a otras áreas de conocimiento. Campbell y Stanley (1995) indican que en estos diseños las unidades de estudio o de muestreo (sujetos) no son asignados al azar a los grupos experimental y control; tienen menor validez interna, ya que pierden control sobre las variables extrañas al no utilizar un muestreo aleatorio.
- **Experimento puro o verdadero.** Aquellos en los que se tienen dos o más grupos de comparación (tratamientos) y un mayor control sobre las variables que influyen, por lo tanto, también un mayor grado de certeza sobre los resultados. Son típicos en investigaciones explicativas (se efectúa manipulación deliberada del factor de estudio por parte del investigador). En ellos se aplican los principios básicos que garantizan una buena experimentación como aleatorización de tratamientos, replicación de tratamientos y el control local del error experimental (incluye al bloqueo).

2. **Según el propósito del experimento** (Concepto, 2025)**.**

- **Determinista.** Se persigue la confirmación de una hipótesis, es decir, se busca demostrar o refutar un principio científico formulado con anterioridad.
- **Aleatorio.** Se desconoce el resultado a obtener, ya que la experimentación simplemente se lleva a cabo para conocer lo que ocurre, es decir, para expandir lo conocido respecto a un tema específico.

3. Según la cantidad de factores de estudio manipulados por el investigador

- **Unifactoriales.** Cuando el investigador manipula de forma deliberada un factor de estudio con dos o más tratamientos (libro DISEÑO EXPERIMENTAL. EXPERIMENTOS UNIFACTORIALES Y PROCEDIMIENTO ESTADÍSTICO EN CIENCIAS AGROPECUARIAS).
- **Factoriales.** Cuando el investigador manipula de forma deliberada y simultánea dos o más factores de estudio (libro DISEÑO EXPERIMENTAL. EXPERIMENTOS FACTORIALES Y PROCEDIMIENTO ESTADÍSTICO EN CIENCIAS AGROPECUARIAS).

1.4.7 Etapas o fases para el establecimiento de un diseño experimental

En una investigación donde se estudian relaciones de causalidad y el investigador manipula un factor de estudio e interviene en las unidades de estudio o de muestreo se debe establecer un experimento y desarrollar un adecuado procedimiento estadístico con fundamento matemático-estadístico que posibilite obtener confiabilidad en los resultados

obtenidos. lo cual constituye la base para poder generar información de calidad a través de la comprobación de hipótesis científicas.

Para que un estudio experimental sea exitoso es necesario realizar diferentes actividades, las cuales se realizan por etapas o fases (descritas por Douglas Montgomery en 2007) el e involucran elementos que se articulan con la metodología de la investigación (subfases 1a y 1b), diseño experimental (subfases 1c, 1d, 1e y 1f), Estadística (fase 2) y finalmente, la escritura científica y la interpretación de los resultados del fenómeno biológico objeto de estudio (fases 3 y 4).

Planificación y ejecución del experimento (involucra seis subfases).

a) **Identificación, enunciado, argumentación, justificación y delimitación del problema de investigación u objeto de estudio.** Requiere que el investigador esté bien informado acerca de los hechos y datos relativos al problema que se expone. Cuanto más claro se presente el problema, mayor probabilidad habrá de solución.
b) **Formulación de hipótesis y declaración de los objetivos del estudio.**
c) **Determinación del o los factores de estudio y sus tratamientos** (puede ser un factor de estudio con dos o más tratamientos (experimentos unifactoriales), o dos o más factores de estudio con dos o más tratamientos cada uno (experimentos factoriales), de acuerdo a su influencia sobre la(s) variable(s) de estudio o respuesta medida(s) en las unidades de estudio o de muestreo (pueden ser plantas, animales, agua, suelo, etc.).

d) **Selección de la(s) variable(s) de estudio a medir** (altura de planta, peso de camarón, producción diaria de leche por vaca, etc.) **en las unidades de estudio o de muestreo.**

e) **Selección del diseño experimental adecuado** (resultados óptimos y ahorro de recursos económicos), de acuerdo a los factores manipulados (un solo factor de estudio en un experimento unifactorial) y a las condiciones del material y entorno experimental (pueden ser un Diseño Completamente al Azar, Diseño en Bloques Completamente al Azar, Diseño Cuadrado Latino simple, Diseño Cuadrado Latino múltiple, Diseño Rectángulo Latino o un Diseño Cuadrado Greco Latino.

f) **Gestión y organización** (de forma previa se debe disponer de todos los recursos necesarios para el desarrollo del experimento), **establecimiento** (se inicia con el replanteo del experimento en el área destinada para ello; ya sea en un laboratorio, invernadero, agroecosistema, hato ganadero, camaronera, etc.) **y conducción del experimento** (todo lo relacionado con el manejo del ensayo el cual debe ser homogéneo en todas las unidades experimentales). Incluye la recolección de datos (se debe considerar la periodicidad de las mediciones (pueden ser de forma diaria, semanal, decenal, quincenal, mensual) y la utilización de instrumentos de medición (los mecánicos que miden magnitudes físicas o documentales que miden dimensiones lógicas) que garanticen la obtención de datos válidos, precisos y exactos.

Procesamiento estadístico de datos (aplicación de métodos estadísticos univariados, bivariados o multivariados de acuerdo al

diseño utilizado). Creación y ordenación de la matriz de datos en software estadístico (es estática, o sea, no cambia en el tiempo).

Análisis, interpretación y discusión de resultados (se analizan los resultados obtenidos considerando los hallazgos a partir del procedimiento estadístico utilizado, se realiza la interpretación del fenómeno biológico estudiado y se comparan dichos resultados con la literatura escrita y que se relaciona con el tema).

Conclusiones y recomendaciones finales

Las conclusiones responden a los objetivos propuestos. Las recomendaciones se establecen de forma opcional y se enfocan en algún estudio pendiente dentro de la línea de investigación.

1.4.8 Conceptos básicos en diseño experimental

Para el desarrollo de una investigación de tipo experimental se necesita el dominio de un conjunto de conceptos básicos relacionados con la metodología de la investigación, diseño experimental y Estadística (en el caso de que exista intervención deliberada del investigador en la manipulación del factor de estudio y las unidades de estudio o de muestreo).

Experimento puro o verdadero

Es la reproducción de algún fenómeno natural en condiciones artificialmente creadas por el investigador.

El objetivo de un experimento es medir en las unidades de estudio o de muestreo el efecto del cambio, que puede provocar el factor de estudio

(causa de variación) en una o varias características estudiadas (variable respuesta o resultado) y minimizar la influencia de variables no controladas que podrían afectar los resultados (Concepto, 2025).

Diseño experimental (DEXPE)

El diseño experimental se refiere a la planificación y estructuración de experimentos de manera explícita y metódica utilizado para medir (recolección de datos) el efecto de una o más variables independientes (factor o factores de estudio) sobre una o varias variables dependientes (variable respuesta). Este enfoque permite obtener resultados válidos y fiables mediante el control de factores externos que puedan influir en el resultado del experimento.

Se define como una estructura de investigación metodológica y estadística planificada, donde el investigador manipula deliberadamente uno o más factores de estudio, controla de forma homogénea el desarrollo del ensayo, bloquea el efecto de factores no controlados (si es posible y necesario), mide el efecto que tienen en otra u otras variables de interés (variable de estudio); para responder a las preguntas de investigación planteadas (hipótesis), encaminadas a cumplir con los objetivos propuestos en el estudio.

El DEXPE es la secuencia completa de pasos previstos de antemano por el investigador para asegurar la obtención de datos apropiados que permitan un análisis objetivo que conduzca a deducciones válidas con respecto al problema en estudio.

Factores presentes en un experimento

En un experimento unifactorial se presentan varios factores, entre los que se encuentran el factor de estudio (constituye la causa de variación en el fenómeno estudiado), los factores controlados (manejo del experimento que debe ser homogéneo) y los factores no controlados (deben ser medidos), que influyen en la variable de estudio o efecto que se mide en las unidades de estudio o de muestreo (Figura 1.1.1).

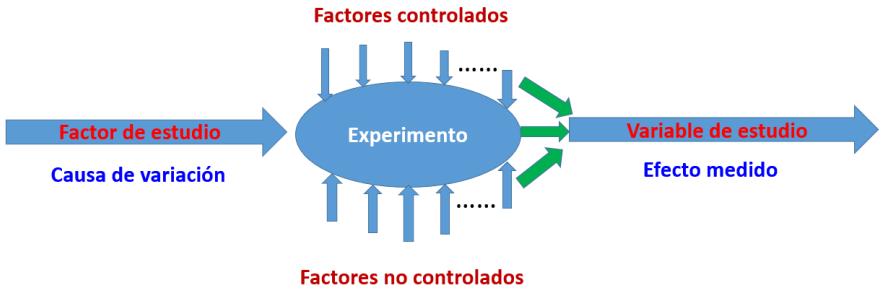


Figura 3.Factores que intervienen en un experimento verdadero.

Factor de estudio (FE)

Es un elemento, material o cualquier procedimiento sometido a estudio o ensayos de comparación sobre el cual el investigador tiene conocimiento y es de su interés evaluarlo. El FE es la posible causa de variación en el fenómeno objeto de estudio, que el investigador manipula de forma deliberada en función de responder a la hipótesis planteada y los objetivos propuestos. En un experimento el investigador puede manipular un FE (se le denomina experimento unifactorial) o dos o más FE (se le denomina experimento factorial). En un gráfico cartesiano el FE se ubica en el eje X.

En la investigación se le denomina causa a la variable que se demuestra que afecta, condiciona o determina la presencia de otra variable que se le llama efecto o consecuencia. La variable causa (factor de estudio) supuestamente modifica la variable efecto, medida en la unidad de estudio o de muestreo (estudios de causalidad). En un estudio experimental el investigador interviene y manipula el factor de estudio, el cual, puede tomar diferentes niveles o versiones (tratamientos). En un estudio observacional se observa la evolución del fenómeno objeto de estudio en condiciones naturales.

El FE se encuentra conformado por niveles o versiones (tratamientos) que pueden ser fijos o aleatorios) en función de la hipótesis que desea comprobar.

Ejemplo. Diferentes tipos de fertilizantes (urea, sulfato de potasio, nitrato de potasio, muriato de potasio) en un experimento agrícola.

Algunos factores de estudio en las ciencias agropecuarias

- Tipo de sustrato orgánico utilizado.
- Sistema de riego.
- Tipos de poda.
- Distancia de siembra
- Densidad de población.
- Temperatura del agua.
- Tipo de alimentación de animales.
- Recambio de agua en una piscina.
- Métodos de control de arvenses.
- Dosis de un probiótico.

Factores controlados (FC)

Los FC son aquellos que se relacionan con el manejo o conducción del experimento por parte del investigador, el cual debe ser homogéneo en cada unidad experimental que conforma el experimento, debido a que se pueden generar sesgos en los resultados.

Ejemplo. Control de arvenses, riego de agua, fertilización, control de plagas, etc., en un experimento sobre densidad de población en el cultivo de zanahoria.

Factores no controlados (FNC)

Los FNC por el investigador son variables que no se pueden controlar durante el desarrollo del experimento, sin embargo, se deben medir, ya que pueden causar confusión en la interpretación de los resultados. Se ubica en el eje Z.

Ejemplos: variables climáticas (temperatura, humedad relativa, precipitación, luminosidad), propiedades del suelo (hídricas, físicas, químicas y biológicas), partículas en el ambiente, niveles de ruido, etc.).

Tratamientos experimentales

Conjunto de circunstancias creadas para el desarrollo del experimento en respuesta a la hipótesis de investigación y que constituyen el centro del mismo. Es una de las categorías del FE (variable explicativa, variable causa). Condición experimental a la cual se someten las unidades experimentales y las unidades de estudio o de muestreo. Es un determinado nivel (cuantitativo) o versión (cualitativo) que forman parte

del FE y es asignado a cualquier unidad experimental de forma independiente. Es lo que se necesita comparar para responder a la hipótesis planteada por el investigador.

Niveles o versiones

Son los distintos valores o cualidades del FE manipulado por el investigador. Para FE cuantitativos, cada valor elegido es un nivel; en este caso el tratamiento del FE se mide y se manipula. Para FE cualitativos, cada cualidad elegida es una versión; en este caso el tratamiento del FE se manipula solamente.

Tabla 1. Ejemplos.

Factor de estudio	Tratamientos	Clasificación
Dosis de un fertilizante	40, 60, 80, 100 kg ha^{-1} de N.	Nivel
Cultivares de soya	INIAP 307, INIAP 308, INIAP 310.	Versión
Dosis de un medicamento	10, 20, 30, 40 mg.	Nivel
Orden del suelo	Andisol, Inceptisol, Oxisol.	Versión
Dosis de un balanceado nutricional	3, 5, 7 kg.	Nivel
Tipo de sustrato orgánico	Humus, compost, estiércol.	Versión
Altura de la poda	5, 10, 15, 20, 25 cm	Nivel

Tratamiento control (testigo)

- Todo experimento verdadero debe contar con al menos un tratamiento testigo.
- Es un tratamiento más dentro del DEXPE.
- Es un punto de referencia necesario, ya que se utiliza para evaluar el efecto de los tratamientos experimentales.

- Existen diversas circunstancias en las que el tratamiento control es útil, debido a que revela las condiciones reales en las que se efectuó el experimento.
- En algunas situaciones es necesario incluir dos tipos de testigo o control distintos. El testigo absoluto (no se aplica ningún tratamiento) y el testigo relativo (se aplica un tratamiento relacionado con el objeto de estudio del experimento).

1.4.9 Ejemplo de un experimento con testigo absoluto y con testigo relativo

Experimento sobre nutrición en caña de azúcar donde se evalúa el efecto de las dosis de N en el rendimiento agrícola del cultivo de caña de azúcar, para ello, se estudian diez tratamientos (dos testigos y ocho tratamientos experimentales donde se aplica las dosis de N, además de un fondo fijo de P_2O_5 y K_2O), ellos son:

1. Testigo absoluto (0 N-0 P_2O_5-0 K_2O), donde no se efectúa aplicación de ningún tipo de fertilizante. Su finalidad es conocer la potencialidad del recurso natural para mantener una producción en el tiempo, debido a que el cultivo de caña de azúcar puede autogenerar fertilidad en el suelo a partir de todos los residuos que produce la planta).
2. Testigo relativo (0 N-40 kg/ha de P_2O_5-0 kg/ha de K_2O), en el cual no se aplica N, aunque si el fondo fijo indicado.
3. 25 kg/ha de N, 40 kg/ha de P_2O_5 y 80 kg/ha de K_2O.
4. 50 kg/ha de N, 40 kg/ha de P_2O_5 y 80 kg/ha de K_2O.
5. 75 kg/ha de N, 40 kg/ha de P_2O_5 y 80 kg/ha de K_2O.
6. 100 kg/ha de N, 40 kg/ha de P_2O_5 y 80 kg/ha de K_2O.

7. 125 kg/ha de N, 40 kg/ha de P_2O_5 y 80 kg/ha de K_2O.
8. 150 kg/ha de N, 40 kg/ha de P_2O_5 y 80 kg/ha de K_2O.
9. 175 kg/ha de N, 40 kg/ha de P_2O_5 y 80 kg/ha de K_2O.
10. 200 kg/ha de N, 40 kg/ha de P_2O_5 y 80 kg/ha de K_2O.

Es importante señalar que en experimentos con animales donde se prueban diferentes tipos de alimentos en condiciones de manipulación y control no puede utilizarse testigo absoluto, ya que por razones obvias no deben tenerse animales sin ofrecerle sus alimentos.

Unidad experimental (UE)

Las UE son los elementos sobre los cuales un tratamiento puede ser asignado de forma independiente. En ellas se establecen o se ubican las unidades de estudio o de muestreo. Una vez expuesta al tratamiento, la UE constituye una sola réplica del tratamiento. Deben ser homogéneas e independientes.

Ejemplos

- Parcelas (donde se aplican diferentes dosis de un fertilizante orgánico a un suelo sembrado con soya), macetas, bolsas de polietileno, placas Petri, etc.
- Peceras (donde se aplican diferentes tipos de balanceado nutricional en camarones), piscinas, tanques, recipientes, etc.
- Galpón (donde se aplican diferentes tipos de medicamento a pollos Broilers), corrales, potreros, jaulas, etc.

Unidad de estudio o muestral (UM)

Elemento que conforma la muestra y sobre el cual se efectúan las mediciones de la(s) variable(s) respuesta(s), o sea, donde se realiza el proceso de recolección de datos. Son los objetos (plantas, animales, suelo, agua, etc.) o sujetos (productores, pescadores, representantes de mascotas, etc.) que realmente se miden, con la utilización de instrumentos de medición mecánicos y/o documentales) y que inicialmente en el experimento deben ser homogéneos en cada unidad experimental.

Ejemplos

1. Si en un experimento la parcela es la UE, una planta, la hoja o los frutos podrían ser las unidades de estudio o muéstrales.
2. Si en un experimento la pecera es la UE, el agua residual o camarón a estudiar podrían ser la unidad de estudio o de muestreo.
3. Si en un experimento el galpón es la UE, los pollos Broilers constituyen la unidad de estudio o muestral.

Variable de estudio

Es la variable respuesta, o sea, el efecto medido en las unidades de estudio o de muestreo. A través de la medición de la variable respuesta se conoce el efecto o los resultados de cada prueba experimental realizada. Su variabilidad se encuentra condicionada por el FE y por las variables intervinientes. Es la variable que se mide mediante la utilización de instrumentos mecánicos (miden magnitudes físicas) o documentales (miden dimensiones lógicas) y que responden a los objetivos del estudio planificados. Siempre se ubica en el eje Y.

Ejemplos. Altura de la planta (cm), peso del animal (kg), materia orgánica del suelo (%), etc.

Error aleatorio (error natural)

Es la variabilidad observada con respecto al valor verdadero, que no se puede explicar por los factores estudiados y que se debe a múltiples factores, generalmente no identificados. Variabilidad no explicada por el modelo. Refleja fluctuaciones naturales. Resulta del pequeño efecto de los factores no estudiados y del error experimental. Aunque siempre existe se debe lograr que sea lo más pequeño posible. Su reducción se logra mediante control experimental y buena aleatorización.

Error experimental (error cometido por el investigador)

Es la variabilidad existente entre observaciones dentro de unidades experimentales tratadas de forma similar, o sea, que se encuentran bajo el efecto del mismo tratamiento (se pueden presentar sesgos de medición al no utilizar instrumentos que posibiliten obtener la máxima precisión y exactitud; así como, sesgos de selección, al no aplicar una técnica de muestreo de forma adecuada), además del efecto que pueda producir el azar.

Error sistemático que altera los resultados del experimento, causado por fallas en la aleatorización de tratamientos, medición o selección de unidades experimentales o unidades de estudio o de muestreo. Componente del error aleatorio que refleja los errores que comente el experimentador en la planeación y ejecución del experimento.

1.4.10 Orígenes del error experimental

1. La variación natural entre UE.
2. La variabilidad en la medición de la respuesta.
3. La imposibilidad de reproducir las condiciones del tratamiento con exactitud de una UE a otra.
4. La interacción de los tratamientos con las UE.

Conclusiones

1. El investigador descubre problemas, construye líneas de investigación y plantea soluciones; solamente si recorre los diferentes niveles de la investigación científica.
2. El diseño de experimentos es un conjunto de técnicas estadísticas y de ingeniería que permiten entender mejor de presentarse situaciones complejas de relación causa-efecto.
3. Las etapas de un diseño experimental comprenden la planeación y ejecución del experimento, el procesamiento de datos, interpretación y discusión de resultados, así como, las conclusiones y recomendaciones finales.
4. El conocimiento e interpretación de los conceptos básicos en la experimentación agrícola constituyen elementos claves como punto de partida en cualquier proceso investigativo.
5. El tipo de diseño de experimentos que se utilice en investigación científica depende de la cantidad de factores de estudio que el investigador manipule y las condiciones del material y entorno experimental.

1.5 Práctica 1. Introducción al manejo de Software estadístico Jamovi

1.5.1 Contenidos

Introducción al manejo del software estadístico Jamovi. Instalación. Barra de menú. Barra de herramientas. Variables, datos, análisis y editar. Visor de resultados.

1.5.2 Objetivos de aprendizaje

Familiarizar al estudiante con el entorno del software estadístico Jamovi y sus funciones básicas (barra de menú, barra de herramientas y visor de resultados) para el procesamiento estadístico de datos en el diseño experimental.

1.5.3 Introducción

En la actualidad el procesamiento de datos se desarrolla con el empleo de ordenadores. Es por ello; que el énfasis se centra en la *interpretación de los métodos de procesamiento de datos* y los resultados que se obtienen; y no en los procedimientos de cálculo utilizados (algoritmos matemáticos previamente establecidos y validados por expertos).

La Estadística es una herramienta utilizada en la investigación científica, teniendo como soporte diferentes programas estadísticos (que realizan la mayoría de las funciones de la Estadística descriptiva, inferencial y representación gráfica de los resultados con el uso y aplicación de distintas técnicas estadísticas), entre los que se encuentran:

1. Microsoft Excel (Análisis básico con hojas de cálculo). Utilizado para cargar matrices de datos.
2. Minitab (herramienta estadística industrial y educativa).
3. R commander (interfaz gráfica para el software R).
4. RStudio (entorno integrado para programación en R).
5. STATA (análisis estadístico aplicado y econométrico).
6. Statistic (análisis estadístico avanzado con interfaz gráfica).
7. Statgraphics (estadística aplicada orientada a calidad).
8. SPSS (Statistical Package for the Social Sciences). Análisis estadístico para ciencias sociales. Tiene una versión de prueba por un mes.
9. PSPP (alternativa libre compatible con SPSS).
10. InfoStat (estadística aplicada a ciencias agrarias, cuenta con una versión gratuita para estudiantes por tres años).
11. Jamovi (estadística moderna con interfaz intuitiva). Software libre y de fácil manejo. Jamovi utiliza R como motor de cálculo, lo que garantiza precisión y flexibilidad; sin embargo, no es necesario programar.

1.5.4 Desarrollo

Jamovi es una Interfaz Gráfica de Usuario (GUI en inglés) de acceso abierto que puede llevar a cabo una amplia gama de análisis estadístico con una serie de menús desplegables y cuadros de diálogos. Permite acceder a muchas capacidades del entorno estadístico R, sin que el usuario tenga que conocer el lenguaje de comandos propio de este entorno.

Puede adquirir datos de casi cualquier tipo de archivo y utilizarlos para generar informes tabulares, gráficos y diagramas de distribuciones y tendencias, estadísticos descriptivos y análisis estadísticos complejos.

1.5.5 Ventajas

1. Fácil de utilizar.
2. Permite el acceso a las funciones y gráficos estadísticos más comunes.
3. Facilita la realización de tareas más complejas.
4. Es multisistema y multiplataforma.
5. Versión demo que permite trabajar directamente desde la nube.
6. Es fácilmente extensible y personalizable.
7. Ofrece la posibilidad de incluir código R.

1.5.6 Procedimiento para iniciar la ejecución del programa Jamovi.

1. Instalación del programa mediante el ingreso al siguiente link https://www.jamovi.org/download.html
2. El usuario debe hacer clic encima del cuadro de diálogo en azul en Jamovi Desktop, descargar la versión de escritorio para Windows y posteriormente se instala la versión disponible más actualizada (2.6.44 Solid), que es la recomendada para la mayoría de los usuarios.



Figura 4. Cuadro de diálogo en azul.

3. Una vez terminada la instalación se procede a anclar en la barra de inicio el programa para mayor facilidad de uso. Se recomienda anclar el programa estadístico Jamovi en la base de tareas del menú inicio.

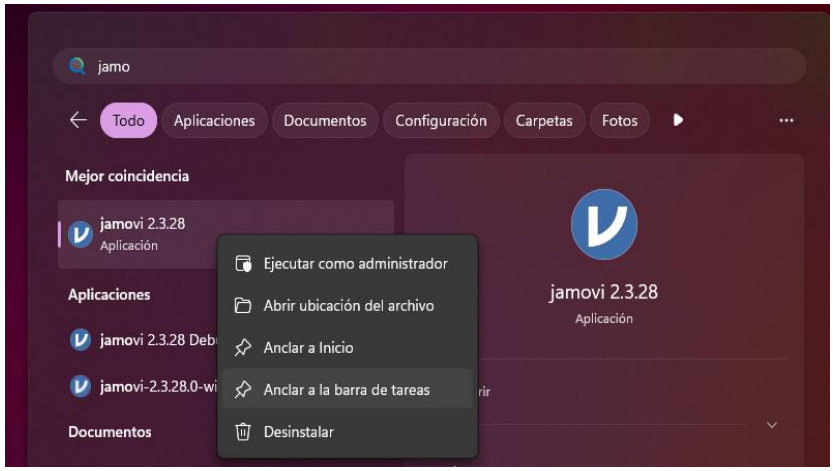


Figura 5. Menú inicio.

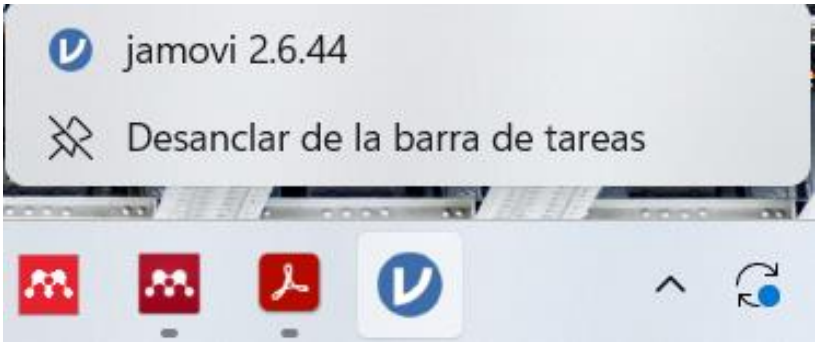


Figura 6. Archivos abiertos recientemente.

Para acceder a los archivos abiertos recientemente haga clic derecho en el icono del programa y después que abra la pantalla de Jamovi-Untitled (archivo sin título), haga clic en la parte superior izquierda ☰ y seleccione de forma directa el archivo que necesita abrir en la parte inferior izquierda en la opción Reciente.

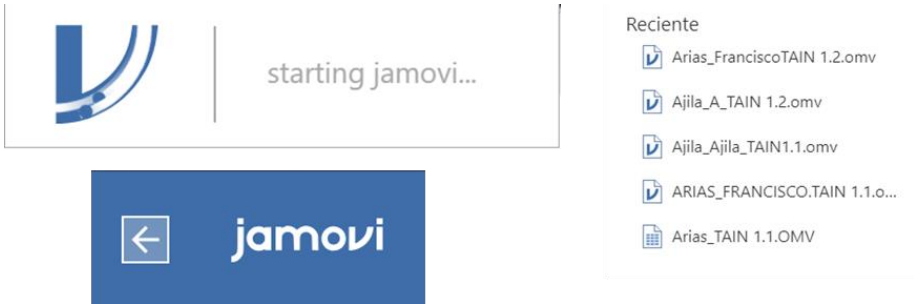


Figura 7. La pantalla de Jamovi-Untitled.

1.5.7 Cuadros de diálogos iniciales

1. Cuadro de diálogo de bienvenida

Haga clic en la parte superior izquierda del Jamovi y se despliega un cuadro de dialogo donde se pueden seleccionar las siguientes opciones:

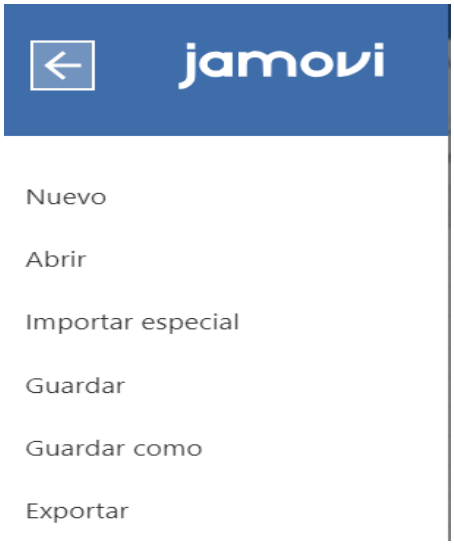


Figura 8.Cuadro de dialogo.

Nuevo: opción para abrir nuevas pantallas de Jamovi, según el requerimiento.

Abrir: utilizada para abrir archivos que se encuentran dentro de tu ordenador.

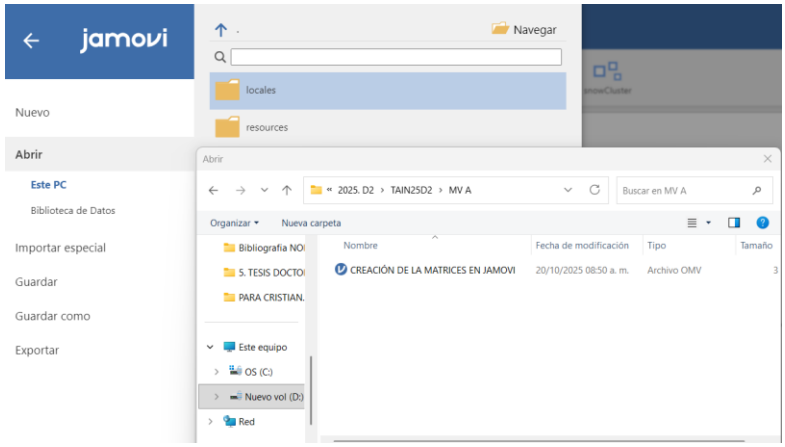


Figura 9.Utilizada para abrir archivos que se encuentran dentro de tu ordenador.

Importar especial: opción que permite importar datos procedentes de archivos creados con otros softwares estadísticos y que se encuentran dentro de tu ordenador.

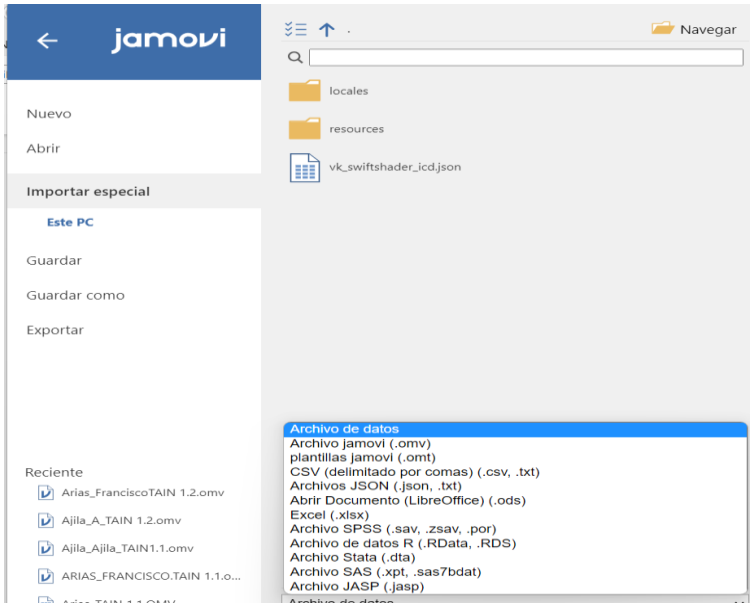


Figura 10.Se utiliza para exportar el archivo.

Guardar: utilizada para guardar lo que obtienes a medida que avanzas en tu procedimiento. Es importante su utilización debido a que puedes perder lo que has obtenido en caso se cerrar el archivo y no guardar, o en caso, de que el ordenador no responde.

Guardar como: se utiliza para renombrar el archivo con un nombre que te permita identificarlo, guardarlo y además de ubicarlo en el ordenador en la carpeta de tu elección, a partir de la opción de Navegar.

Exportar: se utiliza para exportar el archivo a partir de la extensión nativa del software .omv; a otra extensión diferente a la de Jamovi.

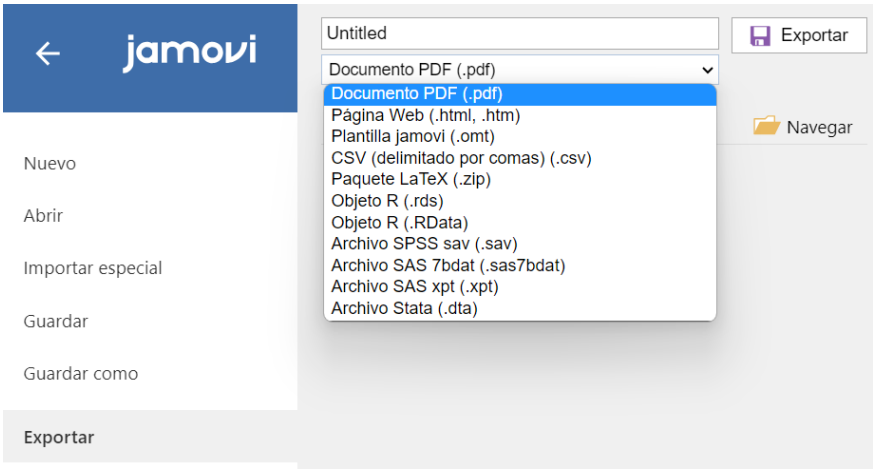


Figura 11.Extensión nativa del software. omv.

2. Variables

Al hacer clic en la pestaña variables; se despliega una hoja donde de forma automática se presentan tres variables, identificadas como A, B y C; y donde se muestran varias opciones. Por defecto, siempre se encuentra, al menos, una variable marcada en el recuadro, pudiéndose

marcar todas las demás y al hacer clic derecho se despliega un cuadro de diálogo con las opciones de agregar una nueva variable, eliminar una existente, además de otras opciones de configuración (editar el nombre de la variable y su descripción) y transformar (se describe en otra sección).

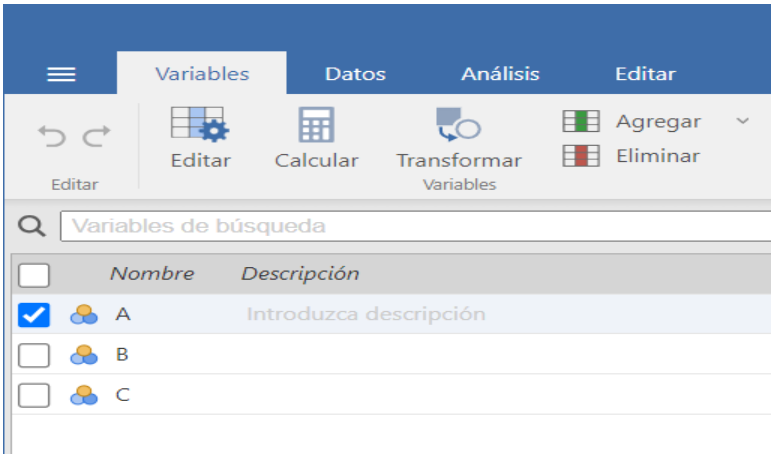


Figura 12.Pestaña de variables.

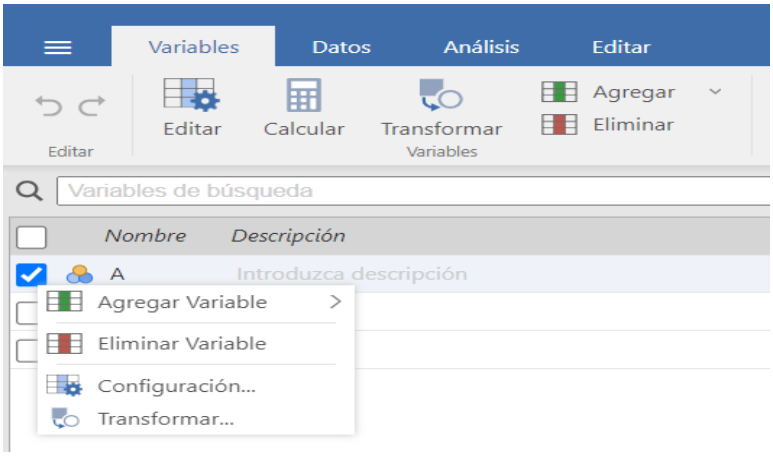
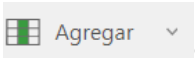


Figura 13.Cuadro de diálogo.

Editar. Se puede editar el nombre de la variable (en la práctica 2 se explicará de forma exhaustiva las posibilidades que ofrece el software en esta opción).

Calcular. mediante esta opción se puede calcular y obtener una nueva variable a partir de los datos de otras variables originales (en la práctica 2 se explicará de forma exhaustiva las posibilidades que ofrece el software en esta opción).

Transformar. mediante esta opción se puede transformar una variable numérica en categórica, a partir de quitarle los atributos de medición origen, distancia u orden (en la práctica 2 se explicará de forma exhaustiva las opciones que ofrece el software en esta opción).

Agregar. al hacer clic en agregar 

o en el signo de adición  se despliega un cuadro de diálogo que permite insertar o añadir una nueva variable de datos, una variable calculada o una variable transformada.

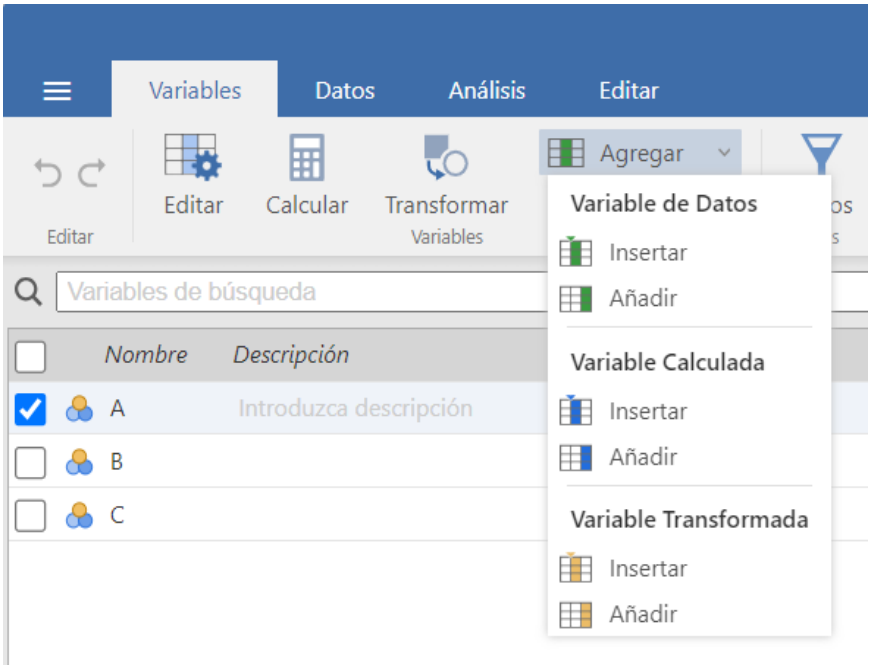


Figura 14.Variable numérica.

Eliminar. Al hacer clic encima de eliminar dará la opción de eliminar la variable o las variables que se encuentren marcada al lado izquierdo.

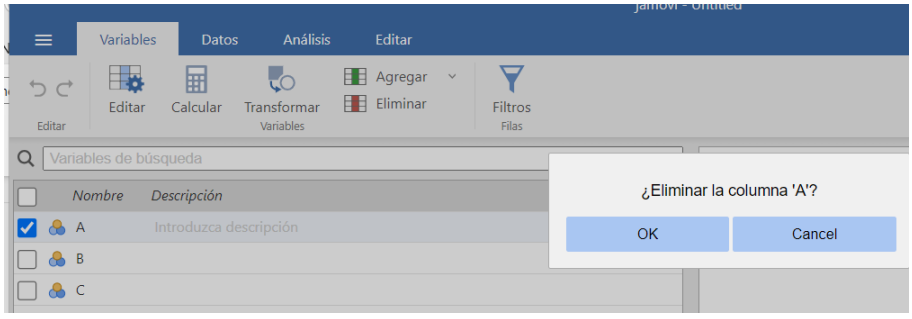


Figura 15.Eliminar la variable.

Filtros. Al hacer clic encima de filtros se despliega un cuadro de dialogo que permite seleccionar solo una parte de los datos con los que quieres trabajar, sin borrar ni modificar el archivo original. Esta opción permite analizar subconjuntos de casos según ciertas condiciones y pueden utilizarse para variables categóricas y numéricas. En la opción f_x se selecciona la variable que se necesita filtrar y posteriormente se coloca la función matemática que se requiere utilizar.

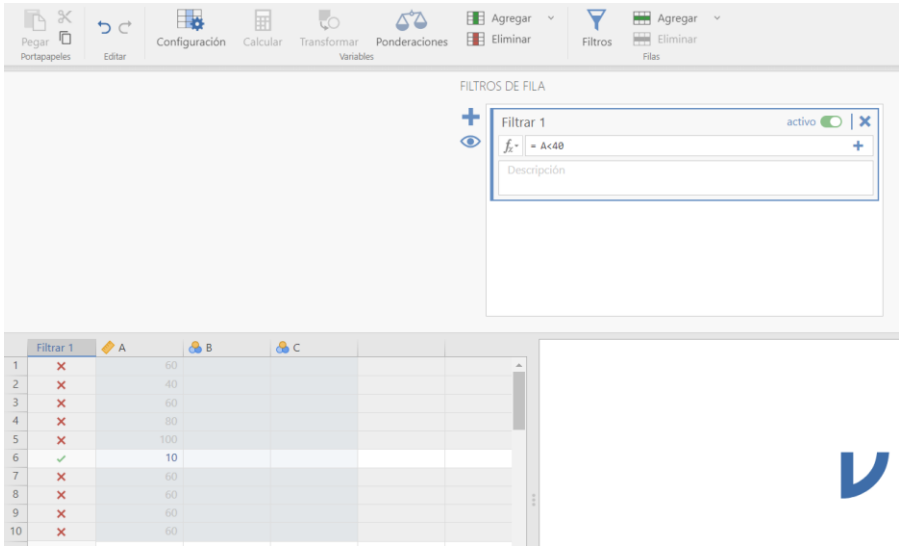


Figura 16.Subconjuntos de casos.

Los filtros permiten excluir temporalmente datos que no cumplen criterios especificados; procesar solo ciertos participantes o casos, comparar grupos de forma más controlada y repetir procesamiento sin volver a limpiar los datos.

Cuando se aplica el filtro los casos que se han excluidos se detallan gris, no son eliminados del archivo, todo el procesamiento estadístico utiliza

solamente los casos filtrados, pudiéndose activar o desactivar el filtro cuando se requiera.

3. Datos

Al hacer clic en la pestaña datos se despliega un cuadro de dialogo con varias opciones, dentro de las que se encuentran pegar portapapeles, editar, para variables aparecen configuración, calcular transformar, ponderaciones, agregar, eliminar, filtros y para las filas se dispone de filtros, agregar y eliminar.

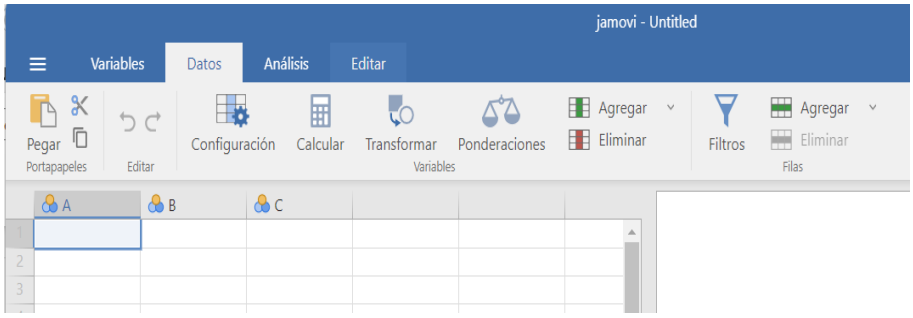


Figura 17.Cuadro de dialogo.

Pegar portapapeles. Se utiliza para copiar y/o pegar información (datos, resultados estadísticos, tablas, gráficos) entre Jamovi y otros programas, como Word, Excel, PowerPoint, Google Docs o correos electrónicos, o sea; permite llevarte lo que se genera en Jamovi a otro archivo sin tener que volver a escribirlo.

Editar. Opciones de deshacer edición (retroceder) o rehacer editar.

Configuración. Al hacer clic se activa la edición y caracterización de las variables.

Calcular y Transformar. Opciones que aparecen también en la sección de variables y que serán desarrolladas en la práctica 2.

Ponderaciones. Se utilizan para dar diferente importancia (peso) a cada caso de la matriz de datos cuando realizas procesamientos estadísticos. Algunos casos pueden contar más o menos según una variable de peso. Se usan cuando la muestra no representa proporcionalmente a la población objeto de estudio, cada caso representa más de una observación, se trabaja con datos de encuestas, se necesita corregir sesgos de muestreo y se necesita obtener estadísticos más reales. La variable de ponderación debe ser numérica, positiva y sin valores perdidos.

Agregar y eliminar. Opciones descritas anteriormente y que se asocian con las variables presentes en la matriz de datos.

Filtros. Descrito anteriormente.

Agregar y eliminar. Opción que posibilitan agregar o eliminar filas o casos.

4. Análisis

Al hacer clic en la pestaña análisis se despliega un cuadro de dialogo con varias opciones, dentro de las que se encuentran diferentes módulos como exploración, pruebas t de Student, análisis de varianza (ANOVA), represión, recuento, factor, esci, modelos lineales, construcción de gráficos, RStudio y el módulo clúster. La mayoría de posibilidades de procesamiento estadístico de datos se utilizarán en las diferentes prácticas que forman parte de este libro.

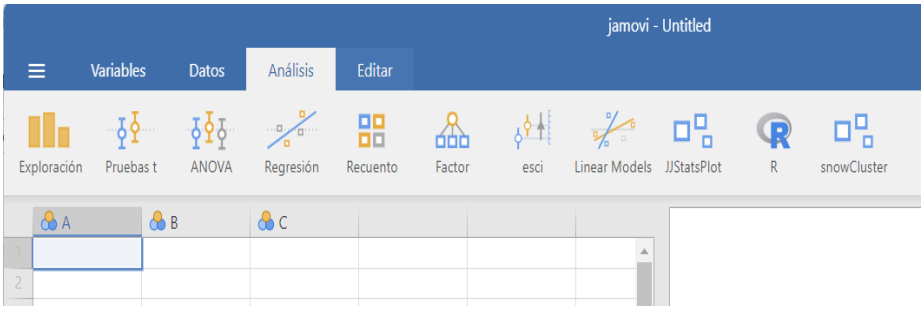


Figura 18.Pestaña análisis.

Exploración. Al hacer clic en la pestaña Exploración se despliegan las siguientes opciones:

- **Descriptivas**. Todas sus opciones se desarrollarán en la práctica 3.
- **Dispersión.** Gráficos de dispersión, gráficos de Pareto. Todas sus opciones se desarrollarán en la práctica 3.
- **Encuestas multivariadas**. Gráfico de encuestas. En Jamovi, la opción Encuestas multivariadas (gráfico de encuestas) se utiliza para visualizar, resumir y comparar simultáneamente las respuestas de múltiples preguntas de una encuesta, especialmente cuando las respuestas son categóricas u ordinales (escalas tipo Likert). Mediante esta opción se pueden reconocer patrones de respuesta entre varias preguntas de una encuesta en un solo gráfico e identificar tendencias de forma rápida y clara, lo cual facilita la interpretación global de los datos



Figura 19.Gráfico de encuestas.

Pruebas t de Student. Todas sus opciones se desarrollarán en la práctica 4.

- Pruebas t de Student para muestras o grupos independientes.
- Pruebas t de Student para muestras pareadas o medidas repetidas.
- Pruebas t de Student para una muestra o grupo.



Figura 20.Pruebas t de Student.

Análisis de varianza (ANOVA). Todas sus opciones se desarrollarán en la práctica 4.

Pruebas paramétricas

- ANOVA de Un Factor.
- ANOVA.
- ANOVA de Medidas Repetidas.
- ANCOVA.
- MANCOVA.

Pruebas estadísticas no Paramétricas

- ANOVA de Un Factor (prueba de Kruskal-Wallis).
- ANOVA de Medidas Repetidas (Prueba de Friedman).

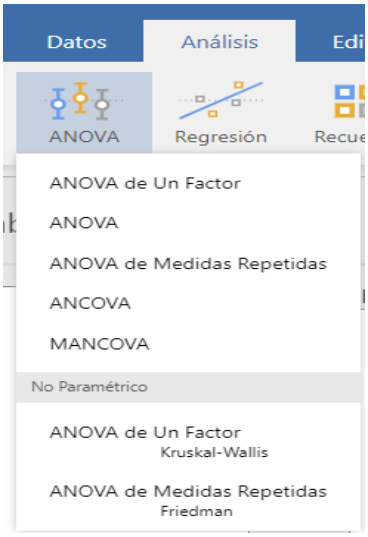


Figura 21.Regresión en Jamovi.

Regresión. Opción que se utiliza para analizar la relación entre una variable dependiente y una o más variables independientes, con el objetivo de explicar, predecir o evaluar efectos entre variables. La opción Regresión en Jamovi se utiliza para analizar, explicar y predecir relaciones entre variables, permitiendo evaluar el efecto de una o más variables independientes sobre una variable dependiente, ya sea continua o categórica.

La regresión permite analizar:

- Coeficientes de regresión (β).
- Significancia estadística (p-valor).
- Coeficiente de determinación (R^2).
- Supuestos del modelo (normalidad de datos, homocedasticidad de varianza, colinealidad).

Matriz de Correlaciones. Opción que se utiliza para analizar de forma simultánea la relación entre varias variables numéricas continuas, mostrando todas las correlaciones posibles en una sola tabla. Permite calcula coeficientes de correlación (Pearson, Spearman o Kendall), muestra la fuerza y dirección de la relación entre las dos variables y permite identificar patrones generales, correlaciones fuertes o posibles problemas de colinealidad.

Correlación Parcial. Se utiliza para analizar la relación entre dos variables controlando el efecto de una o más variables adicionales.

Elimina (controla) la influencia de una variable externa. Muestra la relación pura entre dos variables.

Regresión Lineal. Se utiliza cuando la variable dependiente a predecir es numérica continua.

Regresión Logística. Se usa cuando la variable dependiente es categórica dicotómica (sí/no, 0/1). Permite analizar Odds ratios, probabilidades estimadas y la capacidad predictiva del modelo.

- 2 Resultados (Binomial).
- N Resultados (Multinomial).
- Resultados Ordinales.

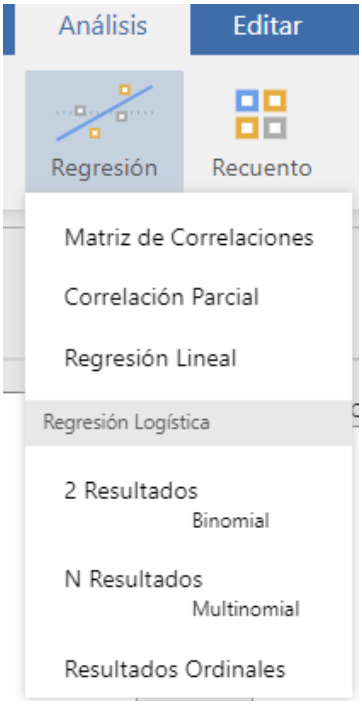


Figura 22.Cuadro de diálogo.

Recuento. Al hacer clic en esta opción se despliega un cuadro de diálogo que se utiliza para contabilizar frecuencias de una o varias variables, especialmente cuando estás trabajando con datos categóricos. Es una herramienta sencilla pero muy útil para explorar la distribución de respuestas antes de aplicar análisis más avanzados. Se utiliza para:

- Conteo de frecuencias absolutas y relativas de variables categóricas.
- Describir de forma rápida la distribución de los datos.
- Detectar valores atípicos o categorías vacías.
- Preparar los datos para procesamientos estadísticos posteriores.

Es una herramienta básica esencial que permite entender la estructura general de los datos antes de avanzar hacia pruebas inferenciales.

Prueba para la proporción en una muestra

- 2 Resultados (Prueba binomial).
- N Resultados (χ^2 de bondad de ajuste).

Tablas de contingencia

- Muestras Independientes (Prueba de independencia de χ^2).
- Muestras Pareadas (Prueba de McNemar).

Regresión Log-Lineal. Es un tipo de modelo estadístico que se utiliza para examinar asociaciones multidimensionales entre variables categóricas, modelar efectos principales e interacciones, analizar tablas de contingencia grandes y complejas; así como, determinar la estructura

de dependencia más adecuada entre variables. El objetivo es entender asociaciones e interacciones entre ellas.

Es la herramienta adecuada cuando se utilizan frecuencias y quieres estudiar relaciones en más de dos variables categóricas simultáneamente.

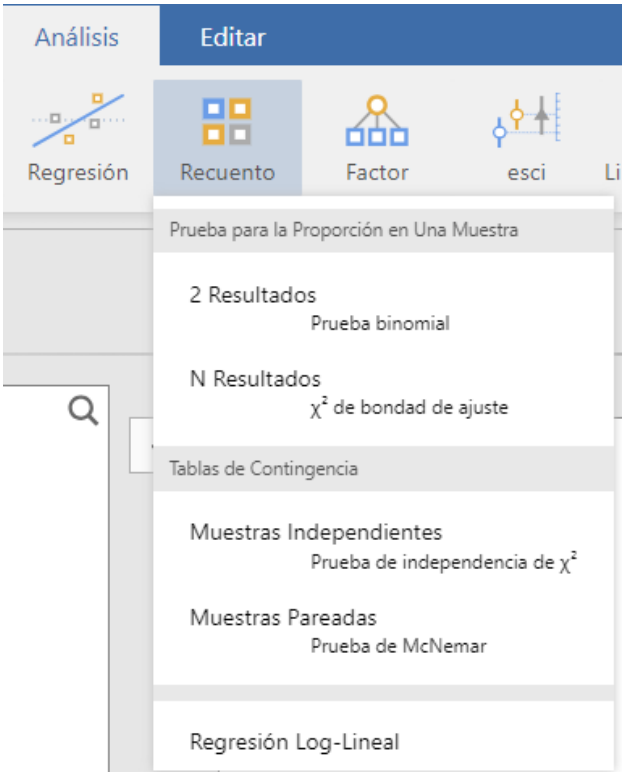


Figura 23.Modelo estadístico.

Factor. Al hacer clic en el módulo Factor se apertura un cuadro de diálogo con las siguientes opciones de procesamiento estadístico multivariado:

Análisis de escala. Análisis de Fiabilidad que se utiliza en la validación de instrumentos de medición documentales.

Reducción de datos. Análisis de Componentes Principales (ACP), que es utilizado para reducir la dimensionalidad de un grupo de variables, Análisis Factorial Exploratorio y Análisis Factorial Confirmatorio.

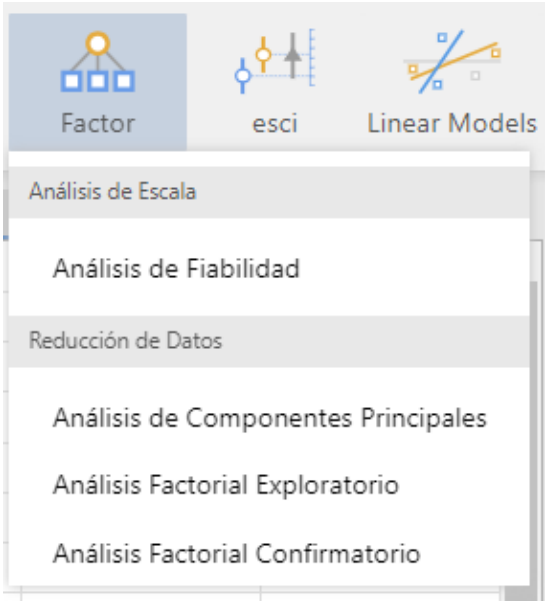


Figura 24.Análisis de Componentes Principales.

Esci. Al hacer clic en el módulo **esci** se apertura un cuadro de diálogo con las siguientes opciones de descripción estadística de variables:

- **Describir.** Calcular estadísticos descriptivos.
- **Medias y medianas** (un solo grupo, dos grupos, pareadas *(mediciones emparejadas/muestras relacionadas)*, contraste entre grupos independientes, factorial 2×2).

- **Correlaciones** (un solo grupo, dos grupos).
- **Proporciones** (un solo grupo, dos grupos, pareadas *(proporciones emparejadas/datos apareados).*
- **Meta-análisis** (medias, diferencia de medias, correlaciones, proporciones, diferencia de proporciones).



Figura 25.Estadística de variables.

Modelos lineales. Al hacer clic en la opción de modelos lineales se apertura un cuadro de diálogo con las opciones GAMLj y GAMLj3, que son módulos del software estadístico orientados al análisis de modelos lineales y mixtos, diferenciándose principalmente en su implementación interna y opciones avanzadas. Esta opción se utilizará a partir de la clase práctica 7 y se despliegan dos módulos con las mismas posibilidades de procesamiento estadístico multivariado.

GAMLj

- Modelo lineal general.
- Modelo mixto.
- Modelos lineales generalizados.
- Modelos mixtos generalizados.

GAMLj3

- Modelo lineal general.
- Modelo lineal mixto.
- Modelo lineal generalizado.
- Modelo mixto generalizado.

El término generalizado indica que el modelo permite utilizar variables dependientes con distribuciones distintas a la normal (binomial, Poisson, etc.).

Los modelos mixtos incorporan efectos fijos y aleatorios, los cuáles son útiles para datos jerárquicos o repetidos.

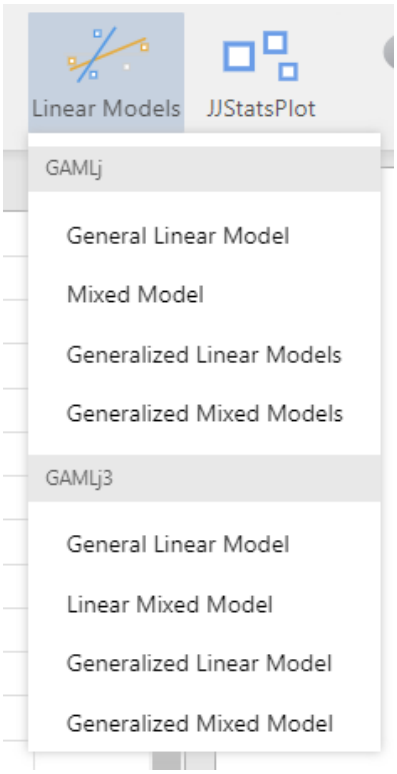


Figura 26.Módulo JJStatsPlot.

Construcción de gráficos. El módulo JJStatsPlot permite construir varios tipos de gráficos univariados y bivariados en función del tipo de variable; o sea, si es continua (numérica) o categórica.

Al hacer clic en el módulo JJStatsPlot se apertura un cuadro de diálogo con las siguientes opciones de representación gráfica de resultados:

1. Numérica continua (histograma). Gráfico univariado.
2. Continua-continua (diagrama de dispersión y matriz de correlación). Gráficos bivariados.
3. Categórica-continua (gráficos Caja-Violín para comparar entre grupos (muestras independientes), gráfico de puntos, gráficos

Caja-Violín para comparar dentro de grupos (mediciones repetidas). Gráficos bivariados.

4. Categórica-categórica (Gráficos de barras (barras simples y barras agrupadas), Gráficos de pastel o circulares. Gráficos bivariados.

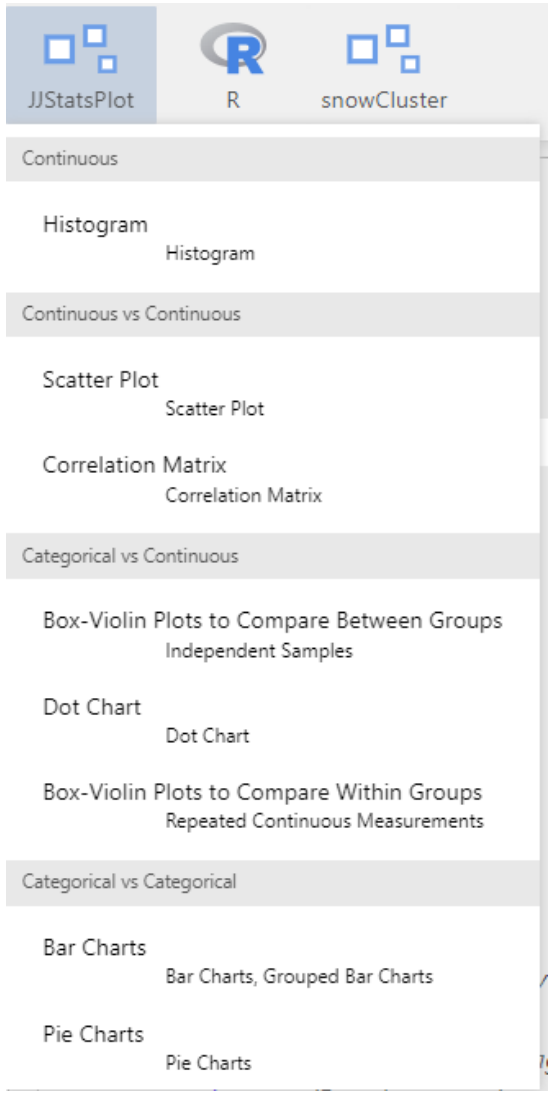


Figura 27.JJStatsPlot.

RStudio. Opción de integración con el software estadístico R.



Figura 28.Software estadístico R.

SnowCluster. módulo que se activa de forma predeterminada cuando se instala el software estadístico, incorpora la función de ARIMA y permite procesar series de tiempo, automatizar la descomposición de datos, evaluar el ajuste del modelo y generar predicciones gráficas de manera sencilla y accesible. Es una herramienta útil para realizar análisis simples de manera accesible.

Al hacer clic en el módulo snowCluster se apertura un cuadro de diálogo con las siguientes opciones:

1. Análisis de clúster (agrupamiento k-medias, agrupamiento jerárquico y agrupamiento de series temporales).
2. Gráficos de clúster (dendrograma de agrupamiento, gráfico de escalamiento multidimensional, análisis de componentes principales y gráfico de grupos).

3. Correspondencia y factor (análisis de correspondencias simple, análisis de correspondencias múltiples, análisis factorial múltiple, análisis factorial de datos mixtos).
4. Clasificación (análisis discriminante lineal, árbol de decisión, aprendizaje automático, análisis o curvas ROC).
5. Series de tiempo (serie temporal univariada, pronóstico con múltiples variables).



Figura 29.SnowCluster módulo que se activa de forma predeterminada.

Editar. Al hacer clic en la barra de menú de despliega la barra de herramientas editar con varias opciones que funciona como un editor de texto para los resultados, entre las que se encuentran:

Pegar (Portapapeles). Pega texto o contenido copiado. Opciones de cortar y copiar.

Editar. Deshacer/Rehacer.

Fuente. Opciones de formato de texto: B. Negrita. I. Cursiva. U. Subrayar. S. Tachar. X_2. Subíndice. X^2. Superíndice. Resaltar color. Color del texto.

Párrafo

- Alinear (izquierda, central, derecha, justificado).
- Lista ordenada. Lista de viñetas. Sangría-1. Sangría+1.

Insertar. Permite añadir elementos al informe: f_x. Fórmula. Inserta expresiones matemáticas (por ejemplo, medias, símbolos estadísticos).

Estilos. Bloque de código. Inserta código (normalmente R o Markdown). Título (H_2). Inserta encabezados para organizar el informe. Enlace. Añade enlaces web o referencias.

Resultados. En esta ventana Jamovi muestra tablas, gráficos, estadísticos y resultados de pruebas

Referencias. Incluye referencias automáticas del proyecto Jamovi, citas de R y paquetes usados; así como, enlaces oficiales (por ejemplo, jamovi.org).

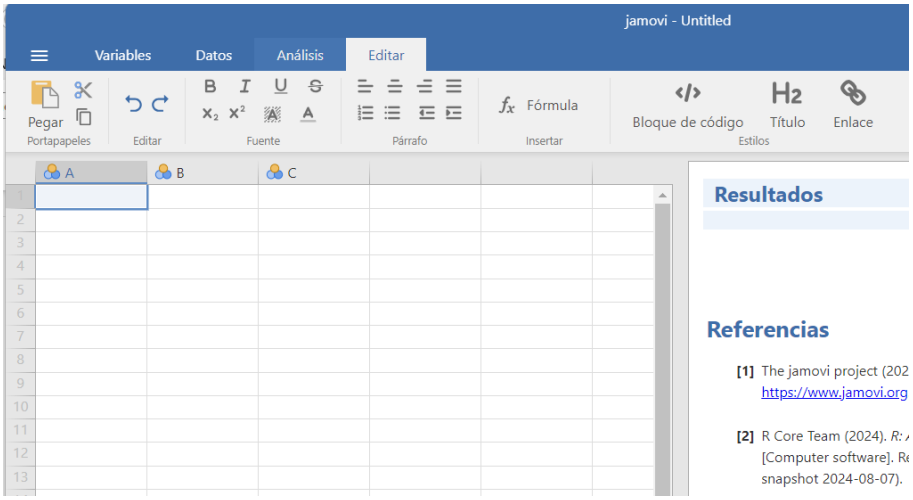


Figura 30.Barra de menú.

1.5.8 Otras opciones de Jamovi

En el panel superior derecho de la pantalla de Jamovi se presentan otras opciones como minimizar (permite ubicar en la barra de tareas del menú inicio el icono del software y mantener de manera activa el archivo), maximizar (permite visualizar otras ventanas de forma simultánea) y cerrar el archivo de Jamovi abierto (se cierra automáticamente si no se ha colocado variables o datos; o te indica si deseas guardar los cambios realizados, cancelar la orden o no guardar los cambios realizados). En la opción que muestra un signo de menor que se puede ocultar la pantalla de datos y solamente se visualizan los resultados del procedimiento estadístico realizado. Si se vuelve a dar clic en se regresa a la opción inicial.

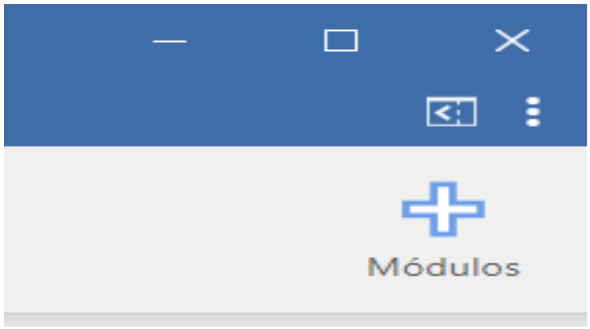


Figura 31.Minimizar.

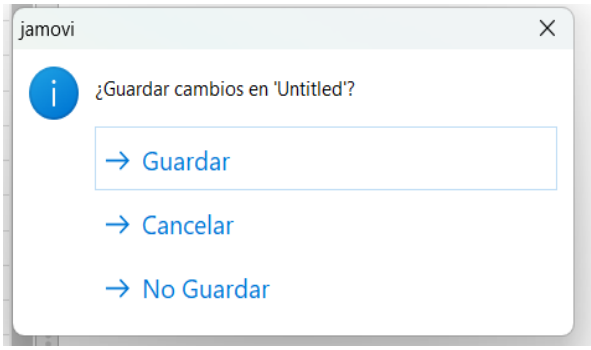


Figura 32.Guardar los cambios.

Al dar clic en los tres punticos ubicados a la derecha de la pantalla se despliegan varias opciones donde se puede aumentar o reducir el tamaño de la hoja (Zoom), establecer en formato numérico y del valor p, las cifras significativas y las cifras decimales, seleccionar las referencias que se ubican o no al final en el visor de resultados, seleccionar modo de sintaxis (describe todos los comandos utilizados). En gráficos; se puede seleccionar el tema del gráfico o la paleta de color, el idioma (varios disponibles y otros en desarrollo) y tipo desarrollador (se describen todos los códigos utilizados ya que se vincula con leguaje de R). En la parte inferior izquierda de la hoja se presentan opciones de

filtrado e información del número de filas, los filtros realizados, datos eliminados y añadidos.

En el panel inferior izquierdo de la pantalla de Jamovi se presentan otras opciones como la cantidad de filtros que se han realizado, el número de filas o casos que tiene la matriz de datos, el filtrado, añadidos y el total de celdas editadas. En el ejemplo son 30 debido a que se dispone de tres variables (A, B y C) y la variable A tiene 10 datos.

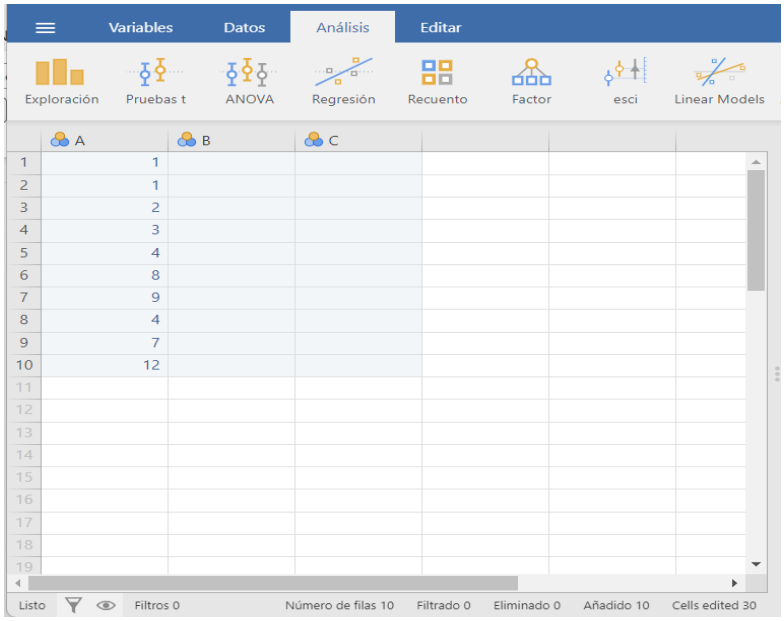


Figura 33.Filtros.

1.5.9 Ejercicio 1. Importar datos de Excel

Contexto

Se dispone de una matriz de datos en Excel con datos de 10 estudiantes de la Facultad de Ciencias sobre edad, sexo, horas de estudio y calificaciones obtenidas en la asignatura de Diseño experimental.

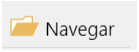
Tabla 2. Importar datos de Excel.

Edad de los estudiantes	Sexo de los estudiantes	Horas de estudio	Calificaciones
21	Masculino	4	9
22	Femenino	3	8
23	Femenino	5	9
24	Femenino	2	6
20	Masculino	1	6
21	Masculino	4	8
23	Femenino	3	8
24	Masculino	4	8
23	Femenino	2	6
25	Femenino	3	7

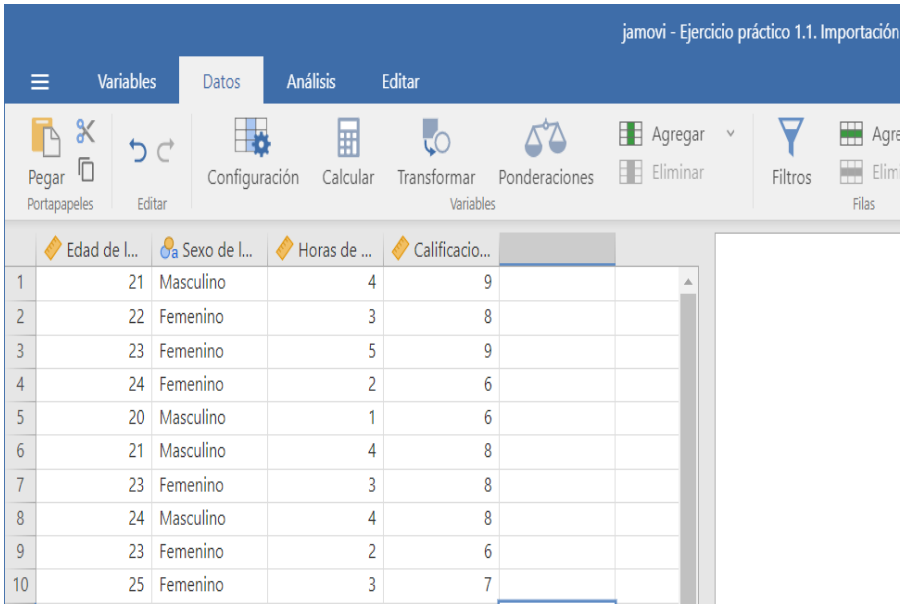
Requerimiento del ejercicio

Importe la matriz de datos creada en Excel.

Procedimiento

1. Abrir el software estadístico Jamovi.
2. Marcar la opción  y seleccionar importar especial.
3. En la parte derecha dar clic en 

para buscar la matriz de datos en Excel.
4. Revisar todas las variables y, si es necesario, editar las características de cada una.



	Edad de l...	Sexo de l...	Horas de ...	Calificacio...
1	21	Masculino	4	9
2	22	Femenino	3	8
3	23	Femenino	5	9
4	24	Femenino	2	6
5	20	Masculino	1	6
6	21	Masculino	4	8
7	23	Femenino	3	8
8	24	Masculino	4	8
9	23	Femenino	2	6
10	25	Femenino	3	7

Figura 34. Software estadístico moderno.

Conclusiones

El desarrollo de la práctica permite la familiarización de manera inicial y funcional con el entorno de Jamovi, un software estadístico moderno, gratuito y de uso intuitivo, ampliamente utilizado en el procesamiento de datos recolectados en la conducción de un diseño experimental. A través del reconocimiento de la barra de menú, barra de herramientas, panel de variables, hoja de datos, opciones de análisis y visor de resultados, el estudiante adquirió una comprensión sólida sobre la estructura y lógica de trabajo de la plataforma. Además, se identificaron las diferencias entre las áreas de ingreso de datos, configuración de variables y visualización de resultados, comprendiendo la importancia de cada una en el proceso analítico. La interacción con el visor de resultados, que actualiza automáticamente las salidas estadísticas,

permitió valorar la eficiencia y transparencia del software en comparación con otros sistemas tradicionales.

La práctica sentó las bases necesarias para que el estudiante pueda gestionar datos, ejecutar estadística básica y organizar resultados, habilidades fundamentales para el desarrollo de competencias en diseño experimental y procesamiento estadístico aplicado. Con ello, se establece un punto de partida esencial para avanzar hacia métodos más complejos en prácticas posteriores.

1.6 Principios y requisitos básicos del diseño experimental

1.6.1 Contenidos

Principios básicos del diseño experimental. Conceptos, objetivos e importancia. Requisitos principales para una buena experimentación. Clasificación de los diseños experimentales.

1.6.2 Objetivos de aprendizaje

1. Reconocer los principios básicos del diseño de experimentos.
2. Identificar los principales requisitos de una buena experimentación.

1.6.3 Introducción

Cuando se realiza una investigación experimental (manipulación del FE e intervención en las unidades de estudio o de muestreo por parte del investigador) y se realiza la implementación (ejecución) del diseño experimental previamente planificado, se requiere de la aplicación de varios principios y requisitos básicos que definen la calidad y precisión

del estudio que se desarrolle; los cuales constituyen elementos clave para el desarrollo de investigaciones válidas y confiables.

El punto de partida para lograr una correcta planeación del diseño experimental es aplicar los principios básicos de aleatorización de tratamientos, replicación de tratamientos y el control local del error experimental (incluye el bloqueo, el cual puede ser sencillo, doble o triple.

El diseño de experimentos trata de fenómenos que son observables (medibles) y repetibles, y donde es importante considerar la ***reproducibilidad*** (poder replicar el experimento, por ello, se debe escribir y explicar de forma detallada todo lo realizado y los recursos empleados) y la ***repetibilidad*** (obtener los mismos resultados).

El resultado de un experimento aporta validez a una teoría; la cual se incrementa cada vez que el experimento es replicado por otros científicos y se obtienen los mismos resultados. La comunidad científica considera la repetición del experimento como el primer requisito para obtener resultados válidos (reproducibilidad) y obtener siempre los mismos resultados (repetibilidad). La ciencia es verificable.

Para que un experimento; pueda catalogarse como un buen experimento, sus conclusiones deben tener validez y precisión. Para cumplir con estas condiciones el experimento de llevarse a cabo teniendo en cuenta los tres principios básicos (aleatorización de tratamientos, replicación de tratamientos y control local de error experimental).

1. Aleatorización de tratamientos

Procedimiento mediante el cual los tratamientos o condiciones experimentales se asignan en las UE (macetas, parcelas, jaulas, placas Petri, peceras, etc.) de forma completamente al azar, con el uso de la tabla de números aleatorios (TNA); de modo tal, que todas las UE consideradas, tengan igual probabilidad de recibir un tratamiento cualquiera. Elimina la relación sistemática entre tratamientos y factores no controlados.

Objetivos o funciones de la aleatorización de tratamientos

- Eliminar sesgos de asignación [evita que el experimentador (consciente o inconscientemente), coloque tratamientos en posiciones favorables o desfavorables]. Garantiza comparaciones justas entre tratamientos.
- Controlar los demás factores o variables denominadas confusoras (distribuye al azar microclima, fertilidad, luz, etc., lográndose que se diluyan sus posibles efectos). Logra mayor validez interna.
- Garantizar que los errores experimentales sean independientes y distribuidos aleatoriamente (garantiza validez en la estimación del error experimental). Permite una mejor precisión y potencia para detectar diferencias reales.
- Evita la influencia de factores no controlados. Garantiza que las diferencias observadas en la respuesta se deban al efecto de los tratamientos y no a sesgos sistemáticos.
- Robustecer los resultados obtenidos a partir de la utilización de la estadística paramétrica. Garantiza inferencias válidas.

Importancia de la aleatorización de los tratamientos

- Aumenta la probabilidad del cumplimiento del supuesto de independencia de los errores (de observaciones, de datos).
- Asegura que las pequeñas diferencias (error aleatorio) provocadas por materiales, equipos y factores no controlados se repartan de manera homogénea en todos los tratamientos.
- Garantiza la validez de los procedimientos estadísticos empleados al realizar comparaciones entre los tratamientos o combinaciones de tratamientos.

2. Replicación o repetición de tratamientos

La replicación constituye uno de los principios fundamentales del diseño experimental al permitir estimar el error experimental y aumentar la precisión de las comparaciones entre tratamientos. Es el número de veces que se aplica un tratamiento, asignados de manera independiente en las unidades experimentales. Replicar o repetir es volver a aplicar un tratamiento, aunque no inmediatamente después de haber corrido el mismo tratamiento; sino, cuando corresponda de acuerdo con la aleatorización de los tratamientos que se realiza de forma simultánea con el uso de la TNA. Aumentar las repeticiones de los tratamientos incrementa la precisión y la capacidad para detectar diferencias reales.

La cantidad de réplicas (repeticiones) necesarias que se deben realizar de cada tratamiento en un experimento no tiene un número correcto, que sea adecuado para todos los casos; dependiendo de factores estadísticos y prácticos.

Recomendaciones generales que se usan en investigación agrícola, biológica y experimental en general para conocer la cantidad de veces que se debe repetir un tratamiento.

1. Un mayor número de repeticiones produce una mayor precisión en las estimaciones que se realicen en el experimento.
2. En investigaciones agropecuarias es frecuente y aceptado utilizar entre tres y cinco repeticiones por tratamiento; sin embargo, no existe un número universalmente aplicable. La cantidad adecuada de repeticiones de cada tratamiento depende, entre otros factores, de la variabilidad esperada del experimento, el tamaño del efecto biológicamente relevante, los recursos humanos y económicos disponibles y la potencia estadística requerida para detectar diferencias entre tratamientos (Gomez y Gomez, 1984; Montgomery, 2013).
3. Sistemas muy variables (campo, plantas, animales) requieren más repeticiones. Sistemas controlados (laboratorios, invernaderos) pueden necesitar menos.
4. El nivel de precisión a obtener. Si se requiere detectar pequeñas diferencias se deben realizar más réplicas de los tratamientos.
5. El costo y la logística necesaria. Si se realizan más replicas se genera mayor costo y trabajo.
6. El incremento de réplicas garantiza el aumento de las observaciones por tratamiento y la utilización de procedimientos estadísticos de datos paramétricos. Lo óptimo es un mínimo de 30 observaciones por tratamiento.

Tabla 3. Replicación o repetición de tratamientos.

Tipo y condición del experimento	Repeticiones recomendadas
Experimentos exploratorios (piloto)	3
Experimentos estándar	4–5
Estudios con alta variabilidad (agronomía, biología)	4–10 repeticiones
Estudios muy costosos o con baja variación	3

Importancia de la replicación de los tratamientos

- Permiten distinguir mejor que parte de la variabilidad total de los datos se debe al error aleatorio y cuál, al factor estudiado, ya que se incrementan los grados de libertad.
- Proporciona la posibilidad de que los efectos de los factores no controlados por el investigador se compensen al reduce su influencia.
- Aumenta la precisión y confiabilidad de los resultados.
- Valida y da soporte científico a las conclusiones del estudio.

Si no se realiza la replicación de tratamientos, un estudio pierde validez y confiabilidad.

3. Control local del error experimental (Incluye la técnica de bloqueo)

Consiste en eliminar o neutralizar cualquier fuente de variación extraña (efectos de cualquier otro factor que pueda intervenir en el experimento) capaz de confundir la acción de los tratamientos del FE.

Su objetivo es lograr cierto grado de aislamiento (control de variables extrañas) que permita al experimentador averiguar con precisión y confiabilidad, cuáles son los efectos, sobre las unidades de estudio, de una (experimento unifactorial) o más (experimento factorial) variables independientes manipuladas por el investigador.

El control local del error experimental incluye a las actividades de manejo que deben ser homogéneas, la medición de factores no controlado y la aplicación de la técnica de bloqueo de ser posible.

El control local describe las acciones que emplea un investigador para:

- Reducir o controlar el error experimental (sesgos de medición y sesgos de selección).
- Incrementar la precisión y exactitud de las observaciones (mediciones con instrumentos mecánicos o documentales certificados).
- Establecer la base de la inferencia estadística (los resultados de la muestra (se calcula el mínimo tamaño muestral con un algoritmo matemático y la aplicación de una técnica de muestreo) se extrapolan a la población de un estudio.

Control de variables

Conjunto de procedimientos físicos, de manejo o estadísticos utilizados para regular el efecto de variables o fenómenos que no son los directamente estudiados. En el trabajo experimental o en los estudios de carácter explicativo debe efectuarse el control de las variables

intervinientes, ya sean factores no controlados (deben ser medidos) y los factores controlados (manejo del experimento, el cual debe ser homogéneo en todo momento y en todas las unidades experimentales.

Técnica de bloqueo

Consiste en distribuir o asignar los tratamientos en las unidades experimentales dentro de cada bloque de forma independiente. Las UE deben ser homogéneas dentro de sí; o sea, dentro del bloque y heterogéneas entre sí; o sea, diferentes entre bloques.

La técnica de bloqueo se utiliza para:

- Controlar la variabilidad natural que existe en experimentos desarrollados en campo (inevitable) y lograr incrementar la precisión entre tratamientos.
- Reducir, atenuar o minimizar la variabilidad o error experimental causado por factores no controlables al incluir grupos homogéneos llamados bloques e incrementar la precisión del experimento.
- Separar la variabilidad conocida y permitir que los tratamientos se comparen en condiciones más homogéneas.

El control local incluye:

- Técnica utilizada (la misma para todas las UE).
- Selección de unidades experimentales (homogéneas dentro del bloque y heterogéneas entre bloques).
- Bloqueo (puede ser sencillo, doble o triple).

- Selección del diseño experimental (depende de la cantidad de FE y tratamientos; así como, de las condiciones del material y entorno experimental (MEE).
- Medición de covariables (FNC con niveles, o sea, cuantitativos).

Al aplicar la técnica de bloqueo se supone que el subconjunto de datos que se obtengan dentro de cada bloque debe resultar más homogéneo que el conjunto total de datos.

Cuando se presenta un factor no controlado con variabilidad sistemática natural en una dirección conocida; la técnica de bloqueo se realiza perpendicular al gradiente de variabilidad del factor no controlado).

La técnica de bloqueo puede aplicarse de forma sencilla, doble o triple. Sobre los diseños en bloques se profundizará en el libro "DISEÑO EXPERIMENTAL. EXPERIMENTOS UNIFACTORIALES Y PROCEDIMIENTO ESTADÍSTICO EN CIENCIAS AGROPECUARIAS".

Importancia del control local del EE

Es utilizado para minimizar el efecto de las variables no controladas (FNC) y acentuar el efecto del o de los FE, si son experimentos unifactoriales o factoriales respectivamente.

Ejemplo práctico de aleatorización y replicación de tratamientos a partir de la construcción del croquis del experimento

Se debe partir del hecho de que ambos principios se aplican de forma simultánea en las unidades experimentales.

Contexto experimental

En un experimento desarrollado en la Universidad Técnica de Machala (UTMACH) en el año 2025 se estudia el efecto del tipo de alimento (balanceado, maíz, soya y pescado) en el crecimiento y desarrollo de pollos Broilers ubicados en jaulas con las condiciones adecuadas. Se conoce que el material y entorno experimental es homogéneo y que se realizarán cuatro réplicas de los tratamientos. Se efectuarán mediciones de las variables de estudio de forma semanal y al final del ciclo de desarrollo de los pollos (unidades de estudio).

Para desarrollar el ejercicio (elaboración del croquis del experimento), previamente, se debe realizar un resumen tabular del contexto experimental que permitirá guiarnos en el proceso de planificación (diseño) del experimento.

Resumen del contexto

Tabla 4.Resumen del contexto.

No.	*Elementos a considerar*	*Cantidad*	*Caracterización*
1	Factor de estudio	1	Tipo de alimento
2	Tratamientos	4	balanceado, maíz, soya y pescado
3	Réplicas	4	Cantidad de veces que se repite un tratamiento
4	Unidades experimentales	4x4=16	Jaulas
5	Material y entorno experimental	-	Homogéneo

Pasos del procedimiento para realizar el croquis del experimento

La aleatorización y replicación de tratamiento se realiza de forma simultánea y se debe considerar cinco pasos para completar y construir el croquis del experimento

1. Seleccionar un punto de partida en la tabla números aleatorios (TNA) (Tabla 1.2.1) donde se escoge al azar un renglón (entre los 40) y una columna (entre las 10); y en la intersección renglón-columna donde se encuentran los cinco números se escoge uno como punto de partida, a partir del cual comienza la asignación de los tratamientos, previo establecimiento de la dirección a seguir (hacia arriba, abajo, izquierda, derecha), lo que depende del número encontrado en el recorrido que se realice en esa dirección (debe ser del uno al cuatro ya que el factor de estudio tiene cuatro tratamientos).

Ejemplo: se escogió el **Renglón 6. Columna 10** y en la intersección se encuentra el número **32716;** dentro de este valor se escoge como punto de partida el número 3 y la dirección a seguir es de izquierda a derecha y cuando se llegue al final del renglón se regresa a la izquierda, aunque se sube al renglón 5 y así sucesivamente. Considerar que cuando te encuentres un número que no se encuentre entre el 1 y el 4 no se considera y se salta. Cada tratamiento debe ser replicado el mismo número de veces, por ello, una vez que un tratamiento alcance el número de réplicas establecido previamente, ya no debes considerarlo en el recorrido.

Renglón	Columna									
	1	2	3	4	5	6	7	8	9	10
1	34600	19108	69812	93480	65191	57359	24408	36527	60414	94913
2	79151	13078	01872	84469	83906	06881	22936	49856	97607	04230
3	92494	97825	58734	08516	37704	20133	70505	06395	54808	57036
4	44852	06858	81140	89296	54813	56856	24316	70468	90027	08372
5	97467	69926	51148	73026	43306	89484	33330	19093	80101	48435
6	96207	18877	70523	29690	44458	99242	35456	39595	87653	32716
7	60337	14292	12704	08359	36120	29596	67888	93498	74984	72836
8	04812	88937	96641	22579	73721	31921	35923	14615	40883	03776
9	30697	44518	57792	97046	99380	17005	30846	55406	22689	88659
10	60331	18044	08728	03094	03465	49651	90558	38744	11275	83301
11	18237	87670	02435	72480	99308	66631	17864	56993	98537	72231
12	98035	63712	25899	61025	35983	46596	59199	36711	03279	15780
13	67961	65714	61082	75324	85711	68100	91197	62429	68027	21201
14	70218	24572	67326	26462	87248	17841	87067	78185	42740	57149
15	83363	17664	88351	55077	07062	17763	60613	60318	05146	02800
16	68761	46051	17313	89765	00076	37890	69373	83061	32370	43278
17	67671	08649	76236	27897	17142	49988	96564	96447	51142	19597
18	95378	01544	76192	69697	29253	70416	17232	38553	21685	22376
19	84149	79121	41425	91820	04102	66022	43084	52345	42530	13834
20	95722	26655	74689	06488	39904	89072	54856	41955	54177	23443
21	19752	28685	28588	43556	66010	50637	37566	74944	20588	98308
22	45683	63873	88430	66485	06903	21488	50694	63228	23797	55052
23	99371	57461	20036	00612	19257	63458	57497	08098	74158	72297
24	72580	53039	43441	98578	54184	45921	65127	01318	68949	48418
25	34315	22973	71948	22061	65262	45078	31623	68896	05562	69511
26	68713	40962	66760	59066	51208	26809	54870	28032	73369	85440
27	44238	95669	78727	90871	08582	59089	73503	97694	68497	94423
28	95424	98332	30624	05323	17194	75596	56225	48613	19599	79610
29	49124	66002	32001	79866	31301	48747	93177	34517	05604	90547
30	74466	36981	62140	54336	98307	84174	31450	67320	24019	93067
31	62705	87371	27786	60655	04768	28167	89910	73654	39125	08345
32	79052	64426	81519	48547	52989	10767	44335	37239	39975	01336
33	77042	31204	27133	25775	26464	74715	90253	64489	09105	40317
34	92857	14367	89222	76064	42903	39980	98344	59800	29486	71233
35	38651	89649	14561	27064	20533	91217	63644	53458	40351	58349
36	09378	90060	82564	32916	71102	81222	74533	11621	10693	96972
37	58472	38563	24415	34840	58615	33807	12195	90969	73059	29497
38	42007	40047	43323	68349	91582	39506	77927	89534	87133	02449
39	53059	66560	80929	47058	64544	75657	68385	92748	31013	97111
40	82725	70760	74764	97880	46162	58002	62728	78882	77898	23641

Figura 35.Tabla de números aleatorios.

Fuente: Universidad de Aguas alientes. España (1993).

2. Establecer un número (sirven de guía en la TNA explicado en el paso 1) y una letra del alfabeto latino para cada tratamiento (se utiliza para construir el croquis del experimento a realizar en el +paso 5) que forma parte del factor de estudio.

Tabla 5.Tratamiento.

No.	Letra	Tratamiento que representa
1	A	Balanceado
2	B	Maíz
3	C	Soya
4	D	Pescado

3. Ubicar las unidades experimentales en el área disponible (laboratorio, campo, invernadero, etc.) para el desarrollo del experimento (la cantidad de UE se obtienen multiplicando el número total de tratamientos (4) por la cantidad de réplicas (4) (16 jaulas en este caso), de forma tal que ubiquen con la misma separación y condición (garantizar homogeneidad).

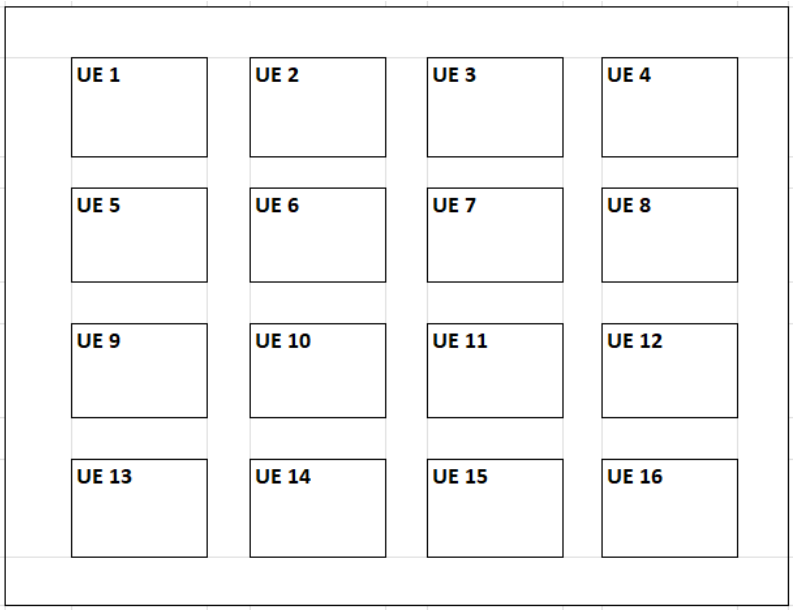


Figura 36.Unidades experimentales.

4. Definir una ruta o camino a seguir en el área experimental para distribuir los tratamientos en las unidades experimentales, de forma tal que, cada UE tenga la misma probabilidad de recibir un tratamiento cualquiera.

Para la asignación de tratamientos en el croquis del experimento se escoge la ruta de izquierda a la derecha, o sea, UE 1-UE 4, UE 5-UE 8, UE-9-UE-12 y UE-13-UE-16.

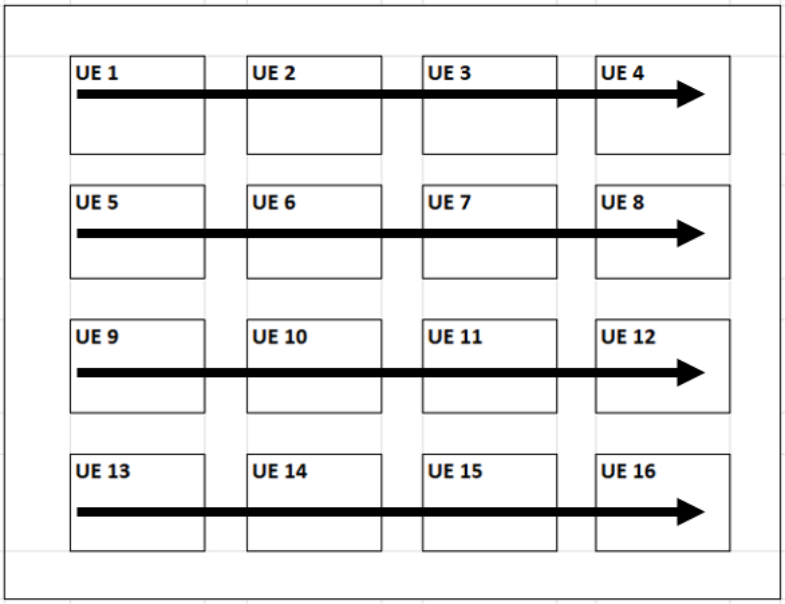


Figura 37.Area experimental.

5. Aplicar el esquema de aleatorización de tratamientos (EAT) correspondiente. En este caso se asignan los tratamientos de forma completamente al azar en las unidades experimentales a nivel de todo el experimento.

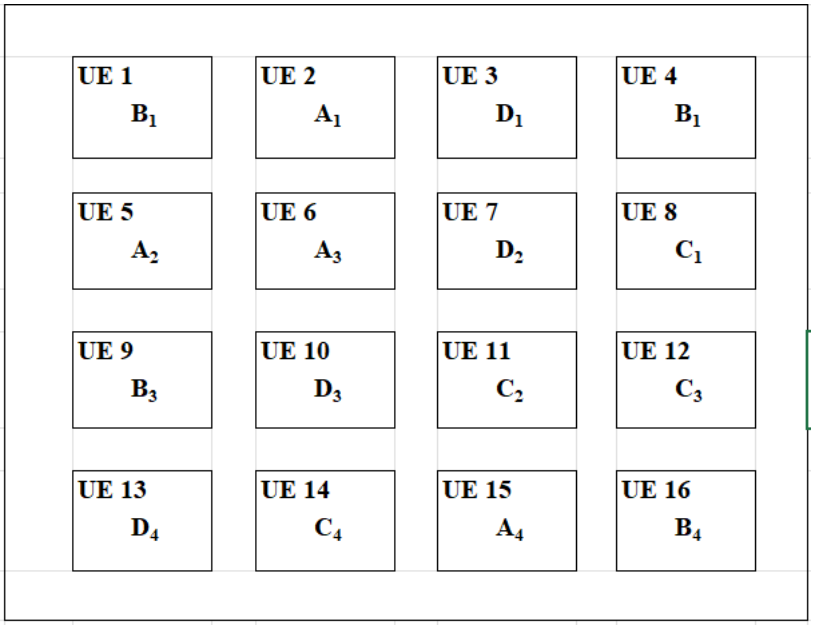


Figura 38.Esquema de aleatorización.

Requisitos principales que debe cumplir un experimento

1. Tipicidad

El experimento se debe desarrollar en condiciones análogas a las del lugar donde se aplicarán los resultados posteriormente (recursos materiales y humanos, tecnología y entorno).

2. Homogeneidad del material y el entorno experimental

La selección del material experimental y el aseguramiento de las condiciones del entorno debe realizarse con cuidado y de acuerdo a los objetivos propuestos (deben ser similares al que se utilizará en la práctica social).

3. Aseguramiento de las condiciones técnico materiales

El experimento científico es costoso, su ejecución debe transitar por etapas ajustadas en el tiempo y la falta de un recurso en el momento

oportuno, puede conducir al fracaso y con ello a la pérdida de todos los recursos invertidos).

4. Garantía de exactitud elevada en el trabajo experimental

En la práctica se pueden cometer errores al iniciar y ejecutar el experimento, y la reducción de errores vulgares (relacionados con la profesionalidad, preparación, disciplina y motivación de los investigadores y auxiliares) y comunes (son inevitables, pero su reducción es posible y necesaria) es tarea de primer orden y en la que no se deben limitar esfuerzos, recursos y medidas organizativas. Cuanto menor sean los errores, mayor será la precisión y exactitud de los datos.

5. Realizar todas las observaciones adicionales del entorno y del desarrollo del experimento

Una documentación amplia y detallada de todo lo ocurrido en el proceso, además, de mediciones adicionales; permite explicar los resultados de forma correcta. Involucra las mediciones del comportamiento del clima y el manejo del experimento.

6. Garantizar la posibilidad de reproducción del experimento

Se relaciona estrechamente con los aspectos anteriores y es el objetivo final a lograr en todo experimento.

Conclusiones

Dentro de los principios básicos para una buena experimentación, la replicación de tratamientos permite estimar la varianza del error experimental, la aleatorización de tratamientos es empleada para obtener una estimación válida de esa varianza y el control local se realiza para eliminar o neutralizar cualquier fuente de variación extraña capaz de confundir la acción de la variable manipulada por el investigador.

Para el desarrollo de una experimentación con calidad se deben cumplir los requisitos estudiados (tipicidad, homogeneidad del material y el entorno experimental, aseguramiento de condiciones técnico materiales, garantía de exactitud elevada, observaciones adicionales y garantizar la posibilidad de reproducción del experimento).

1.7 Práctica 2. Edición y transformación de variables con Jamovi.

1.7.1 Contenidos

Editor de variables. Categorización de variables. Cálculo de nueva variable, recodificación y transformación de variables con el empleo de software estadístico.

1.7.2 Objetivos de aprendizaje

1. Desarrollar habilidades en la edición y transformación de variables dentro de Jamovi para optimizar el procesamiento estadístico de datos en diseño experimental.
2. Editar las características que identifican a las variables que forman parte de la matriz de datos.
3. Calcular una nueva variable a partir de los datos procedentes de otras variables medidas en escala numérica.
4. Recodificar variables a partir de la información emitida por variables originales.

1.7.3 Introducción

El manejo adecuado de las variables es un paso fundamental que asegura un procesamiento estadístico de datos de forma correcta. Por ello, Jamovi ofrece una interfaz intuitiva que permite editar, configurar,

calcular una nueva variable y transformar variables de manera sencilla, lo cual garantiza que los datos sean procesados de forma correcta con el uso del Software y la interpretación de los resultados que realiza el investigador responda a los objetivos planificados.

1.7.4 Desarrollo

El menú de opciones del Jamovi ofrece varias posibilidades, dentro de las cuales se encuentra ***variables*** que se utiliza para definir las características de las variables, como sus nombres, información adicional sobre ellas (descripciones, etc.). También se puede editar, calcular, transformar, añadir, eliminar, así como, filtrar por filas (Pajankar, 2022)**.**

Al aperturar el Software Estadístico Jamovi y hacer clic en el menú variables se despliegan las opciones de editar, calcular, transformar, añadir, eliminar y filtrar variables

Editar. Al hacer clic en la opción editar es posible modificar características de la variable; el nombre (si es continua se le debe colocar la unidad de medida entre paréntesis al final del nombre de la variable), cuidando en todos los casos la ortografía; su descripción (algunas características de la variable que se necesiten adicionar), el tipo de medida (nominal u ordinal si es una variable categórica; escala si es numérica; además de la opción ID, o sea, colocar un número como identificador del caso o unidad de estudio); el tipo de dato (si corresponden al tipo de medida nominal u ordinal se marca texto y se adicionan sus categorías en la opción niveles; si es continua se coloca entero si es discreta, o sea, si se obtiene de un conteo, o decimal, si en

la medición que se efectúa se obtiene al menos un decimal. Se recomienda utilizar todos los decimales que entregue el instrumento de medición mecánico utilizado) (Pajankar, 2022). Además, se presenta la opción de valores perdidos, que se utiliza para marcar, identificar y gestionar datos faltantes; permite que el software ignore valores incorrectos, no elimine información valiosa y produzca análisis más fiables.

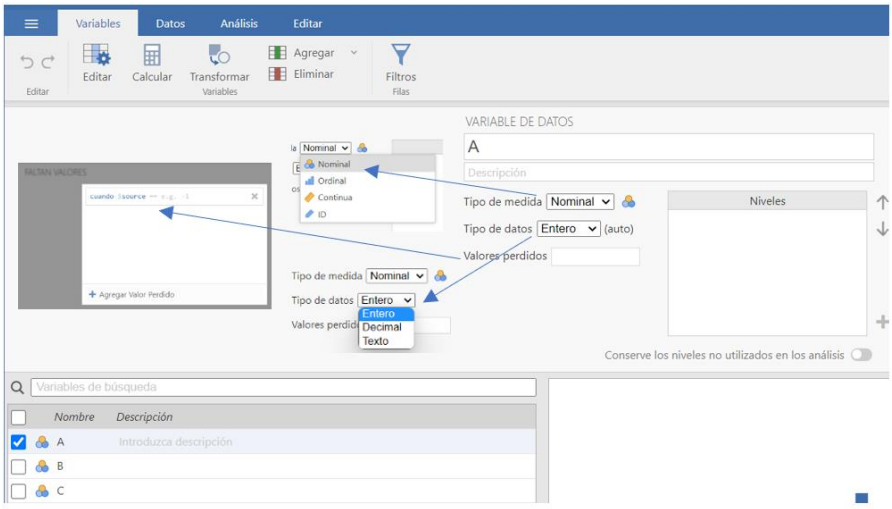


Figura 39.Características de la variable.

Calcular. Al hacer clic en la opción calcular variable de Jamovi posibilita calcular y obtener una nueva variable a partir de los datos obtenidos en otras variables originales.

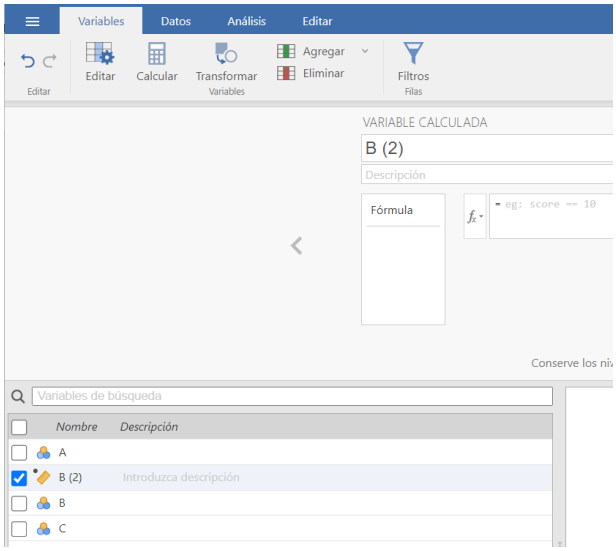


Figura 40.Variable de Jamovi.

La función de la opción calcular; consiste en realizar cálculos mediante ecuaciones matemáticas basadas en operadores aritméticos, lógicos, exponenciales, etc**.**

Las llamadas variables calculadas permiten generar nuevas columnas a partir de fórmulas que involucren otras variables, facilitando operaciones como promedios, sumas, puntuaciones invertidas, transformaciones logarítmicas o estandarizaciones.

1.7.5 Ejercicio 3. Calcular una nueva variable.

Contexto

Se dispone del peso (kg) y la altura (m) de 10 estudiantes de la Facultad de Ciencias Agropecuarias de la Universidad Técnica de Machala.

Tabla 6.Calcular una nueva variable.

Peso (kg)	**58,4**	**68,2**	**78,6**	**58,8**	**70,2**	**66,4**	**68,2**	**62,3**	**81,0**	**59,7**
Altura (m)	1,65	1,54	1,78	1,64	1,69	1,52	1,49	1,70	1,72	1,56

Requerimiento del ejercicio

1. Calcule el Índice de Masa Corporal (IMC) de los estudiantes considerando la fórmula ($\text{IMC} = \frac{\text{Peso (kg)}}{(\text{Altura (m)} * \text{Altura (m)}^2}$).

Procedimiento estadístico

1. Abrir el software estadístico Jamovi.
2. Renombrar el archivo (ejercicio 2.1) y ubicarlo en la carpeta creada en su ordenador para ubicar los ejercicios.
3. Hacer clic en el menú variables; y en la opción editar para ubicar el nombre de cada variable, unidad de medición, tipo de medida y tipo de dato.

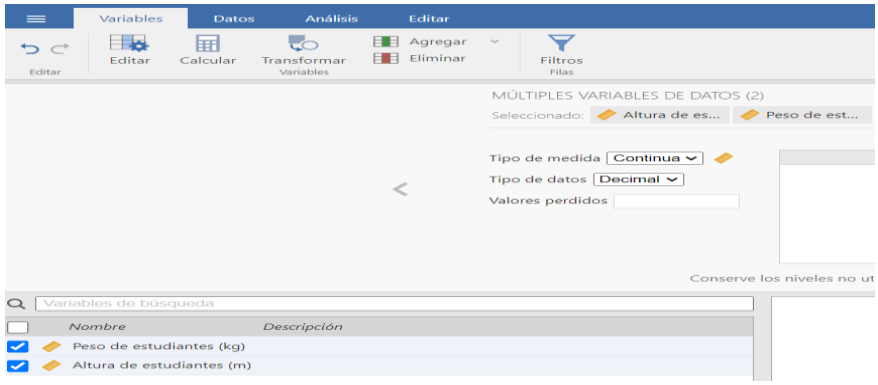


Figura 41.Menú variables.

4. Hacer clic en el menú datos y tipiar los valores correspondientes a las variables peso y altura de estudiantes. Se debe utilizar el punto para los decimales, ya no reconoce la coma.

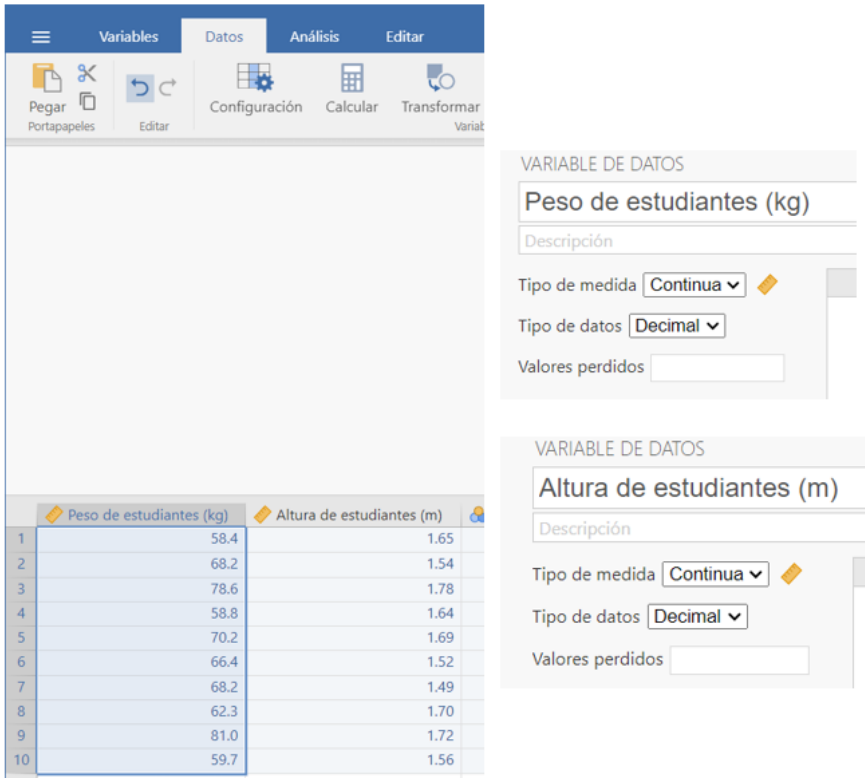


Figura 42.Menú datos.

5. Ubicar el cursor en la tercera columna y hacer clic en calcular, se coloca el nombre y la descripción de la nueva variable calculada y posteriormente se aplica la fórmula que se obtiene en funciones ($IMC = \frac{Peso\ (kg)}{(Altura\ (m) * Altura\ (m)^2}$), se seleccionan las variables originales involucradas al hacer doble clic y finalmente aceptar.

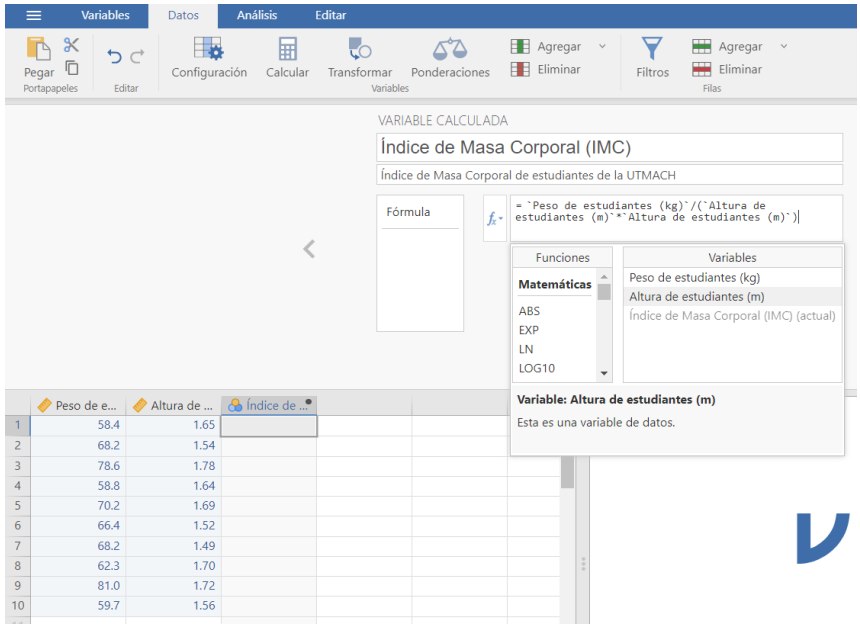


Figura 43. Variables originales.

6. Ubicar el cursor al final de la fórmula y aceptar.

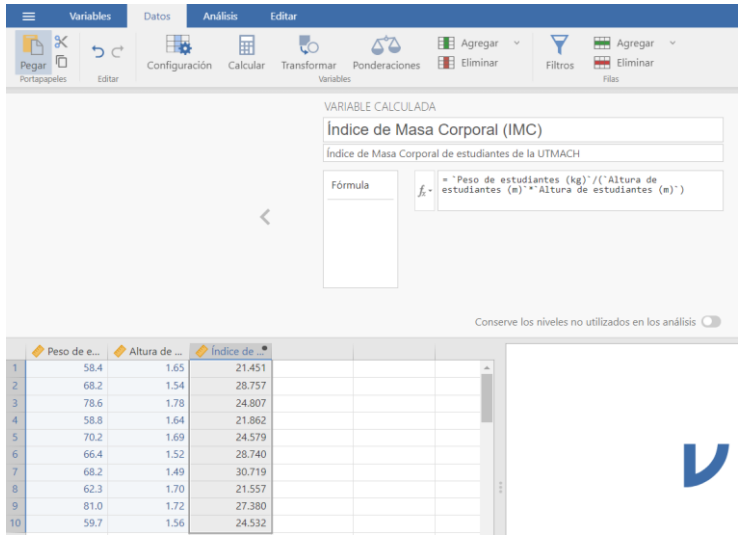


Figura 44. Transformaciones matemáticas o cálculos.

Transformar. La transformación de variables en Jamovi consiste en crear nuevas variables derivadas a partir de operaciones, recodificaciones, transformaciones matemáticas o cálculos aplicados sobre variables existentes. Este proceso es esencial para preparar los datos antes del procesamiento estadístico y para generar indicadores compuestos, puntuaciones, escalas y más.

La manipulación de los datos mediante la transformación de variables permite:

1. Crear nuevas variables que recopilen de forma general la información de varias variables y adaptadas a tus necesidades.
2. Generar nuevas variables que categoricen los datos obtenidos en una escala numérica.
3. Preparar datos para otros procesamientos estadísticos.
4. Facilitar la interpretación de los resultados.
5. Convertir variables numéricas en categóricas. Puede convertir una categórica ordinal en categórica nominal.

Si se cuenta con una variable numérica de razón para transformarla en categórica ordinal se quitan, a la primera, los atributos de medición origen y distancia. Si se cuenta con una variable numérica de razón para transformarla en categórica nominal se le quitan, a la primera, los atributos de medición origen, distancia y orden.

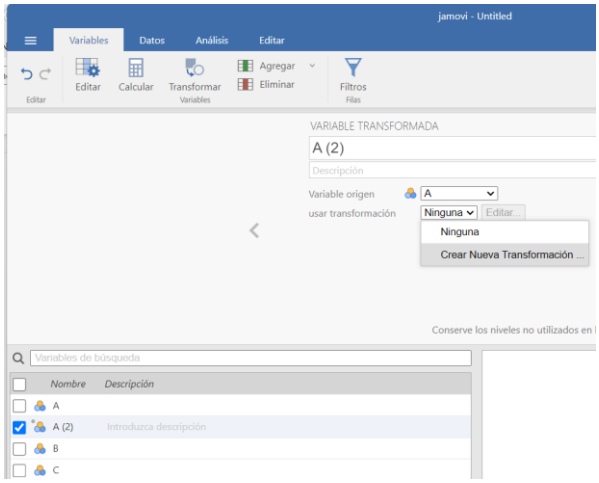


Figura 45.Transformación de variables.

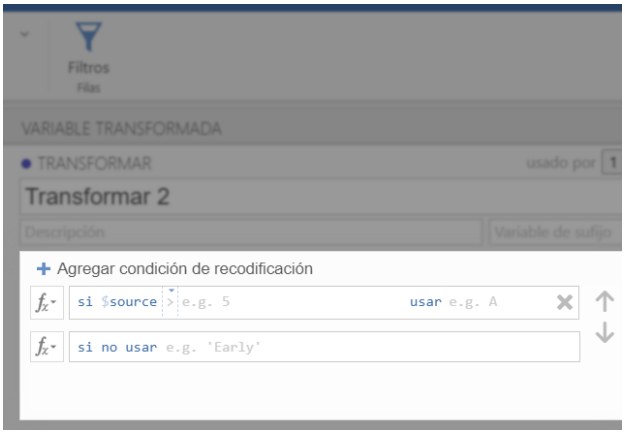


Figura 46.Recodificaciones.

1.7.6 Ejercicio 4. Transformación de variables

Contexto

Se dispone del IMC de 10 estudiantes de la Facultad de Ciencias Agropecuarias de la Universidad Técnica de Machala.

Tabla 7.Transformación de variables.

Estudiantes	1	2	3	4	5	6	7	8	9	10
IMC	21,451	28,757	24,807	21,862	24,579	28,740	30,719	21,557	27,380	24,532

Requerimiento del ejercicio

1. Considere el Índice de Masa Corporal (IMC) obtenido de los estudiantes universitarios y las categorías establecidas por la Organización Mundial de la Salud (OMS), que se muestran a continuación.

Tabla 8.Requerimiento del ejercicio.

Categoría	Rango de IMC (kg/m²)
Bajo peso	< 18,5
Peso normal	18,5 – 24,9
Sobrepeso (Pre-obesidad)	25,0 – 29,9
Obesidad Clase I	30,0 – 34,9
Obesidad Clase II	35,0 – 39,9
Obesidad Clase III)	≥ 40,0

Fuente: OMS (2026).

Procedimiento

1. Abrir la matriz de datos del ejercicio práctico 2.1 con Jamovi.
2. Renombrar el archivo (ejercicio 2.2) y ubicarlo en la carpeta creada en su ordenador para ubicar los ejercicios.

3. Ubicar el cursor en la cuarta columna y hacer clic en la cuarta columna.

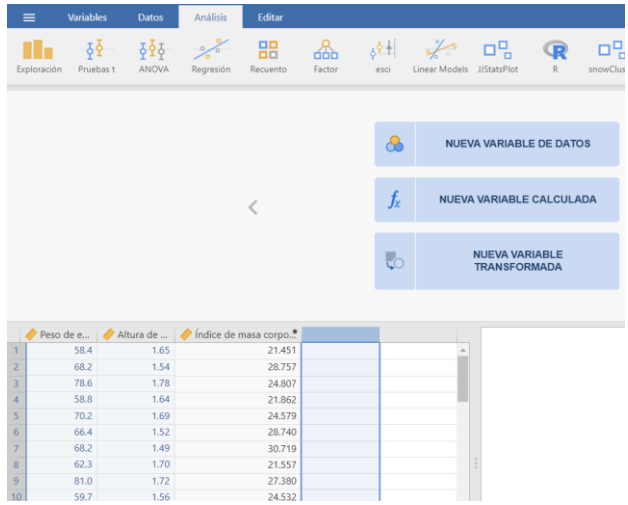


Figura 47.Matriz de datos

4. Hacer clic en la opción nueva variable transformada, renombrarla y seleccionar la variable de origen; o sea, la que utilizará para la transformación. Hacer clic encima de la variable seleccionada, la cual se considera como variable de origen.

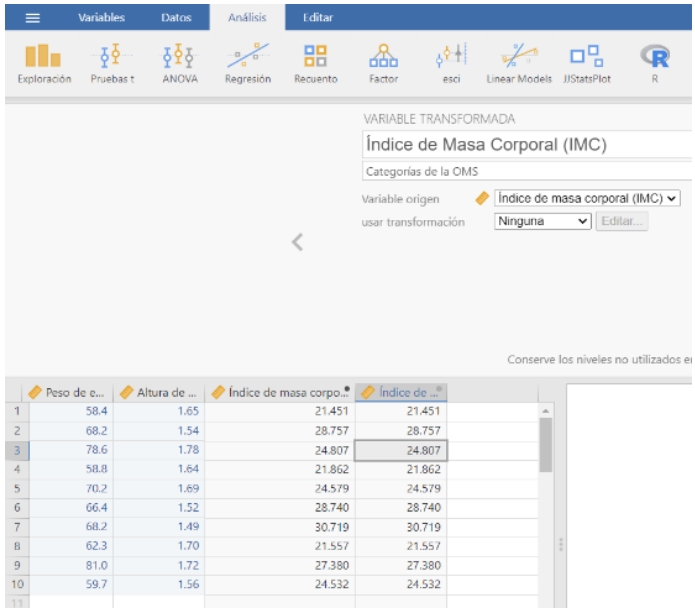


Figura 48.Variable de origen.

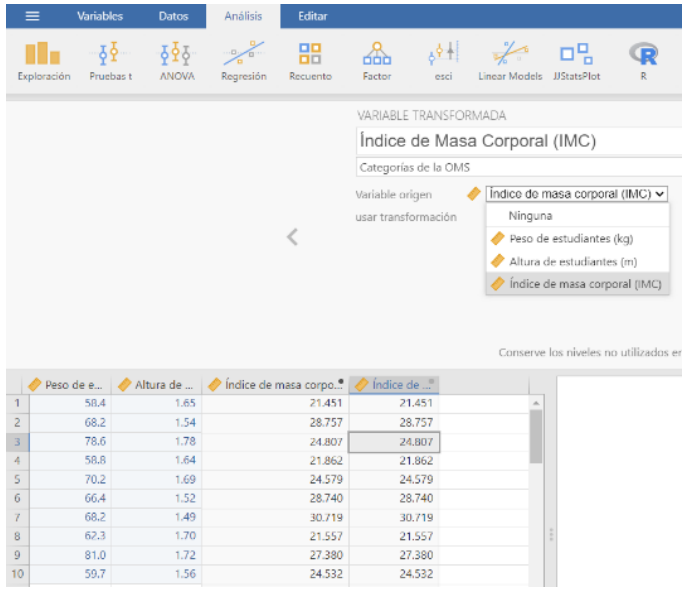


Figura 49.Categoría de IMC

5. Desplegar la viñeta en la opción usar transformación y marcar la opción de Crear Nueva Transformación, donde se debe hacer clic en la opción agregar condición de recodificación cada vez que se agregue una categoría de IMC



Figura 50.Nueva Transformación.



Figura 51.Transformación.

6. Al agregar condición de recodificación se apertura una fila donde indica la función y usar que es la que se debe utilizar para colocar los rangos de las categorías de IMC



Figura 52.Condición de recodificación.

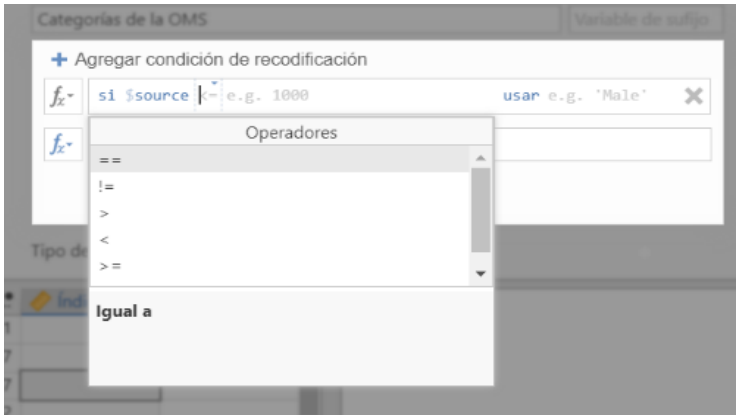


Figura 53.Categorías de IMC.

7. En este cuadro de diálogo se puede escoger una función matemática, escoger en source (fuente), el rango (< 18,5 kg/m^2) de la primera categoría, en este caso Bajo peso. Considere que los decimales se separan por punto, siempre trabajar con el valor extremo del intervalo (por ejemplo, la categoría peso normal su intervalo es 18,5-24,9, por ello, se debe colocar ≤ 24,9) y la categoría se debe colocar entre comillas para que el software la reconozca.



Figura 54.Cuadro de diálogo.

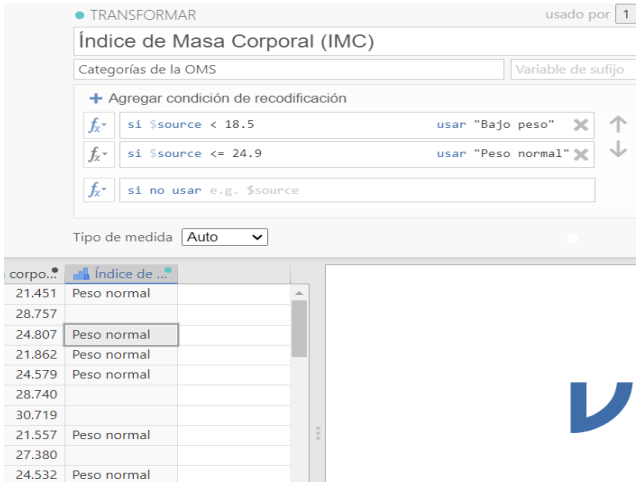


Figura 55.Función matemática.

8. Continuar agregando los rangos y categorías establecidas por la OMS para concluir con el ejercicio.

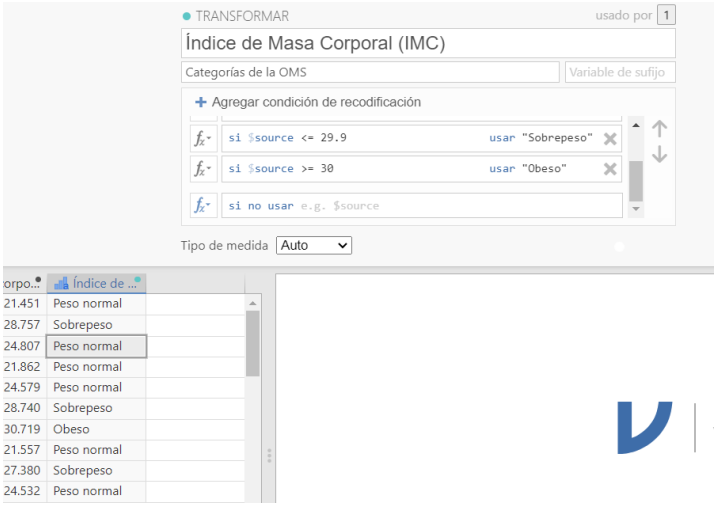


Figura 56.Rangos y categorías.

Conclusiones

La edición, cálculo de nueva variable y transformación de variables son opciones que presenta Jamovi que constituye un componente esencial en el flujo de trabajo de los procesamientos estadísticos, que garantiza que cada variable sea interpretada según su naturaleza real y que pueda adaptarse a las necesidades concretas del tipo de estudio que se desarrolla.

Conocer las funciones proporciona un mayor control sobre el conjunto de datos y permite aprovechar al máximo el potencial del software estadístico Jamovi.

CAPÍTULO II

2 APLICACIONES PRÁCTICAS DE LA ESTADÍSTICA

La Estadística constituye una disciplina fundamental dentro de las ciencias naturales, sociales y aplicadas, ya que ofrece un conjunto de métodos y técnicas que permiten transformar datos en información útil que posibilita la comprensión de fenómenos y la toma de decisiones fundamentadas. La Estadística es una herramienta poderosa que permite comprender y transformar la realidad. En el sector agropecuario es indispensable para enfrentar los retos que plantea la producción de alimentos en un mundo caracterizado por la variabilidad climática, la degradación de ecosistemas y la necesidad de un manejo sostenible de agroecosistemas.

En este capítulo se profundiza en cómo los métodos estadísticos permiten transformar datos provenientes del campo en información valiosa para la gestión eficiente de los sistemas agrícolas y pecuarios. La utilidad de la Estadística en este ámbito se manifiesta en múltiples dimensiones:

- **Producción agrícola:** el procesamiento estadístico de datos de rendimientos de cultivos, la evaluación de fertilizantes y el estudio de plagas permiten identificar patrones y tomar decisiones que incrementen la productividad.
- **Ganadería y producción pecuaria y acuícola:** la Estadística se emplea para evaluar parámetros de crecimiento, reproducción y producción de leche o carne. El uso de modelos de regresión

facilita predecir el comportamiento de los animales en función de factores como la alimentación o el manejo sanitario.

- **Economía agropecuaria:** los estudios estadísticos de precios, costos de producción y tendencias de mercado ayudan a los productores a planificar sus inversiones y reducir riesgos financieros. Las encuestas y muestreos permiten conocer la percepción de los consumidores y orientar la oferta hacia sus necesidades.
- **Medio ambiente y sostenibilidad:** clave en el monitoreo de suelos, agua y biodiversidad.

La Estadística aplicada al sector agropecuario contribuye a mejorar la eficiencia productiva y fortalece la capacidad de los productores para enfrentar la incertidumbre y tomar decisiones basadas en evidencia. Este capítulo busca mostrar que la Estadística es un aliado estratégico en la construcción de sistemas agroalimentarios más competitivos, sostenibles y resilientes; a través de ejemplos concretos se demuestra que la Estadística es un puente entre la teoría y la práctica.

2.1 Estadística descriptiva

2.1.1 Contenidos

La Estadística como ciencia. Historia. Conceptos fundamentales. Variables. Criterios de clasificación. Medidas de resumen de datos. Características, ventajas, desventajas e importancia práctica. Representación gráfica de resultados.

2.1.2 Objetivos de aprendizaje

1. Reconocer los principales conceptos relacionados con la Estadística como ciencia.
2. Identificar los tipos de variables de acuerdo a diferentes criterios de clasificación.
3. Definir los principales elementos teóricos y prácticos relacionados con las medidas de resumen de datos obtenidos en variables numéricas y categóricas; así como, la representación de los resultados obtenidos en investigaciones en el área de las ciencias agropecuarias.

2.1.3 Introducción

La Estadística es una herramienta matemática potente y eficaz que permite arrojar luz sobre cualquier tipo de estudio que, a causa de su variabilidad intrínseca, no pueda ser abordado desde la perspectiva de las leyes deterministas. En la medida en que permite obtener resultados prácticos, es una excelente guía para la acción en situaciones que entrañan incertidumbre.

La Estadística descriptiva es la rama de la Estadística encargada de organizar, clasificar, resumir y presentar datos de manera que puedan interpretarse fácilmente. Su propósito principal es describir el comportamiento de un conjunto o serie de datos de las variables medidas en el estudio. Para ello, emplea tablas (promedios, proporciones, variabilidad) y gráficos (univariados, bivariados y multivariados) que permiten identificar tendencias, dispersión y patrones relevantes. Gracias a estos procedimientos, la Estadística descriptiva facilita la

comprensión de fenómenos biológicos, la interpretación de los resultados y la toma de decisiones fundamentadas en los datos.

Historia breve de la Estadística

En la historia de la humanidad la Estadística ha jugado un rol determinante en el procesamiento de datos obtenidos en las diferentes áreas del conocimiento, dominios académicos y líneas de investigación.

Orígenes en la Antigüedad

Civilizaciones antiguas como Egipto, Babilonia y China utilizaban conteos y registros para fines tributarios, militares y administrativos. Los censos eran fundamentales para conocer la población, calcular impuestos y organizar ejércitos. En el año 3 000 a. C. se tienen evidencias de registros estadísticos relacionados con tierras y producción agrícola.

Edad Media y primeros registros sistemáticos

Durante la Edad Media, los señores feudales realizaban enumeraciones de siervos y tierras para fines económicos y militares. En el siglo IX se consolidaron registros de propiedad y población. A partir del siglo XVII, en Inglaterra, se publicaron los Bills of Mortality (listas de mortalidad), que clasificaban nacimientos y muertes por género y causa.

Nacimiento de la probabilidad (siglo XVI – XVII)

Los juegos de azar impulsaron el desarrollo de la teoría de la probabilidad. Girolamo Cardano fue uno de los primeros en estudiar matemáticamente el azar. Posteriormente, matemáticos como Pascal y

Fermat formalizaron la teoría de la probabilidad, sentando las bases de la Estadística moderna.

Desarrollo en los siglos XVII y XVIII

John Graunt (1662) analizó 30 años de listas de mortalidad en Londres, realizando predicciones sobre nacimientos y muertes, lo que se considera uno de los primeros estudios demográficos. La Estadística comenzó a vincularse con la demografía, la economía y la administración pública.

Consolidación en el siglo XIX

Se desarrollaron métodos más rigurosos de muestreo y análisis. La Estadística se aplicó en ciencias sociales, biología y medicina. El término se consolidó como disciplina científica vinculada al estudio de datos poblacionales y económicos.

Estadística moderna (siglo XX en adelante)

Con el avance de la informática, la Estadística se expandió hacia la Estadística inferencial, el análisis multivariante y la modelización matemática. Se convirtió en una herramienta esencial en investigación científica, economía, ingeniería, salud y ciencias sociales. En la actualidad, la Estadística es clave en el procesamiento de grandes volúmenes de datos (Big Data) y en la inteligencia artificial.

La contribución de la Estadística en la formulación de nuevas teorías ha sido decisiva, sin embargo, el desarrollo actual de nuevos y sofisticados softwares estadísticos que simplifican el procesamiento estadístico

exigen a estudiantes y profesionales una sólida base teórica-práctica; así como, la una actualización constante, la cual debe enfocarse en el conocimiento de los procedimientos estadísticos encaminados a apoyar la toma de las decisiones con fundamento científico.

Estadística

Estadística escrita con mayúscula inicial se refiere a una ciencia o disciplina formal (rama de las matemáticas) basada en metodologías, técnicas y procedimientos utilizados para recopilar, organizar, clasificar, procesar y presentar series o conjuntos de datos. Se divide en Estadística descriptiva e inferencial. Por otro lado, estadísticas (como sustantivo, escrito en minúscula y en plural), se refiere a los datos o resultados numéricos que se obtienen de la medición de variables de estudio que caracterizan a las unidades de estudio o de muestreo.

Conceptualización

La palabra estadística proviene del término alemán que significa "ciencia del estado". Su origen data del año 2 200 a. C. con los censos chinos ordenados por el emperador Tao.

La definición y conceptualización de la Estadística como ciencia es muy amplia y diversa (Existen más de 100 (Hurtado, 2000) y se encuentra asociada con la experiencia de cada profesional en su área del conocimiento.

Concepto actual

Ciencia encargada de desarrollar métodos, técnicas y procedimientos para el manejo de datos, su ordenación, procesamiento, descripción y presentación que contribuyen: al estudio científico de los problemas planteados, a la adquisición de conocimientos sobre las realidades en estudio, a la toma de decisiones y a la mejora de la práctica desarrollada por cualquier tipo de profesional. (Hurtado, 2000).

Concepto enfocado en la investigación

La Estadística le ofrece al profesional en formación, herramientas fundamentales (Flores, 2012), que le permiten desarrollar competencias investigativas, al obtener conocimientos, habilidades y destrezas encaminadas a su utilización en la búsqueda de soluciones a situaciones problemáticas creadas en su entorno (Salcedo, 2013).

Concepto relacionado con las etapas del proceso de investigación

La Estadística es una herramienta necesaria en investigación científica que posibilita aplicar diferentes técnicas y procedimientos en cada momento por el que transita la línea de investigación, encaminados a la recolección, organización y procesamiento estadístico de datos, con la finalidad de resumirlos o realizar inferencias válidas, que posibiliten realizar una adecuada interpretación del fenómeno objeto de estudio con un fundamento matemático y el apoyo a la toma de decisiones (Figura 2.1.1).



Figura 57. Etapas del proceso de investigación científica donde la Estadística se encuentra presente.

2.1.4 Conceptos fundamentales

Datos

Conjunto de observaciones o hechos que, una vez recogidos, organizados y procesados, se transforman en información o conocimiento. Constituyen los valores o atributos obtenidos en el estudio, recolectados mediante la utilización de técnicas e instrumentos de medición (mecánicos o documentales).

Pueden ser:

- **Primarios**: cuando el investigador realiza las mediciones directamente y controla los sesgos de medición. **Ejemplos. Valor:** Altura de la planta de banano (m). **Atributo:** Forma de la hoja (acorazonada, lanceolada, ovalada, etc.)
- **Secundarios**: cuando se obtienen de instituciones, organismos o de otros investigadores, o sea, el investigador no efectúa la

recolección de datos; y, por tanto, no tiene control sobre los sesgos de medición. **Ejemplos. Valor**: Producción mundial de soya obtenido de la página oficial de la FAOSTAT (https://www.fao.org/faostat/es/#data/QC/visualize), **Atributo:** Provincias que forman parte del Ecuador obtenido de la página oficial del Instituto Nacional de Estadística y Censos (https://www.ecuadorencifras.gob.ec/estadisticas/)

Los datos se recolectan con la utilización de instrumentos de medición mecánicos (mide magnitudes físicas) y documentales (mide dimensiones lógicas) (Tabla 9).

Tabla 9. Técnicas de recolección de datos en función del tipo de dato obtenido en el estudio

Tipos de datos	**Unidades de estudio**	
	Objetos	**Sujetos**
	Técnicas de recolección de datos	
Primarios	Observación	Observación, encuesta, entrevista, Psicometría
Secundarios		Documentación

En la actualidad el proceso de manejo de datos se realiza de forma rápida y sencilla gracias al uso de ordenadores y del desarrollo de nuevas tecnologías. El procesamiento estadístico de los datos proporciona información por medio de números o frecuencias, es decir, son datos que han sido organizados para tener un significado útil (información) que sirva de ayuda en la toma de decisiones, o sea, los datos se transforman en información después de su procesamiento estadístico (Tabla 10).

Tabla 10. Transformación de materia prima en productos.

Materia prima	Producto
Petróleo	Gasolina
Madera	Papel
Datos	Información

Población objeto de estudio (POE)

Es el conjunto de unidades de estudio (objetos o sujetos) delimitadas espacial y temporalmente que cumplen con criterios de selección (inclusión y exclusión) y que poseen alguna característica común medible que se desea estudiar. La POE debe ser delimitada en el espacio y en el tiempo (delimitadores de la POE) y cumplir con criterios de selección y a partir de este valor se efectúa el cálculo del tamaño muestral.

Puede ser:

Finita: se conoce el número total de elementos que la componen.

Infinita: no se conoce el número total de unidades de estudio que la conforman.

Parámetro

Dato que describe una característica de interés en la POE. Las medidas de resumen de datos calculadas en la POE se denotan con letras griegas (Ejemplo: media, varianza, desviación típica o estándar de los datos, proporción). El tamaño de la POE se denota por la letra N.

Muestra

Subconjunto de unidades de estudio (debe ser representativa de la POE), utilizada para realizar el procesamiento de datos y extrapolar las conclusiones obtenidas a dicha POE. El tamaño muestral se denota por la letra n.

Estadígrafo o estadístico

Función definida que resume los datos obtenidos en una muestra (siempre es aleatorio). Se denotan con letras latinas (media, varianza, desviación típica o estándar de los datos, error típico o estándar de la media, proporción).

Unidad de estudio o de muestreo

Cada elemento o entidad (objetos o sujetos) que conforma la POE o la muestra, y sobre la cual se efectúan las mediciones de las variables objeto de estudio.

Constante

Características de la POE o de la muestra que no cambian en estado y expresión, por ello, a estos datos no se le realizan cálculos estadísticos, debido a que no se presenta variabilidad.

Variables

El conocimiento acerca de las *variables* constituye una parte esencial en la metodología de la investigación, diseño experimental y Estadística; y

es uno de los elementos más importantes en cualquier proceso investigativo que se desarrolle en los diferentes campos del saber.

Importancia de las variables en investigación científica

1. Del tipo de variable (numérica o categórica) depende el procesamiento estadístico de los datos que se realice.
2. Del tipo de variable (independiente, intervinientes y dependientes) depende la interpretación de los resultados del diseño experimental que se realice por parte del investigador.

Definición

Es una propiedad (valor) o atributo (cualidad) con respecto al cual los elementos de una muestra, o una población objeto de estudio (objetos o sujetos); se diferencian en algo verificable; y cuya posible variación puede ser observada (para ello, se utilizan instrumentos de medición mecánicos o documentales) en las unidades de estudio o de muestreo (Hernández-Sampieri et al., 1991).

De forma resumida se puede definir a una variable como una característica (propiedad o atributo) medible en las unidades de estudio o de muestreo, lo cual constituye su condición indispensable.

Las variables en investigación científica

Gran parte de la investigación en ciencias e ingenierías se realiza con el uso de la experimentación (intervención del investigador en la manipulación del FE y las UM), aunque se desarrollan también estudios no experimentales.

Los métodos estadísticos (paramétricos o no paramétricos) pueden incrementar en gran medida la eficiencia de los experimentos y de la investigación, y fortalecen las conclusiones obtenidas (argumento matemático).

En toda investigación (experimental o no experimental) se observan (medición) fenómenos biológicos en objetos o sujetos (unidades de estudio), para cumplir con ciertos propósitos (hipótesis) investigativos (se efectúan mediciones de variables de interés (variable de estudio) que responden a los objetivos establecidos previamente).

Criterios de clasificación

La clasificación de variables en función de diferentes criterios

Se realiza teniendo en cuenta varios criterios, complementados unos con otros y asociados de manera general a su forma de expresión (Figura 2.1.2) (Rodríguez et al., 2018). En este libro se detallarán los elementos relacionados con los puntos 3, 4, 5 y 6.

1. Objetividad.
2. Dimensionalidad.
3. Naturaleza de los datos.
4. Escalas de medición.
5. Relación causa-efecto.
6. Amplitud de las mediciones.
7. Niveles de la investigación.



Figura 58.Criterios de clasificación de las variables y donde se utilizan.

En la taxonomía de las variables se debe partir de:

- Parsimonia (realizar la clasificación lo más sencillo posible).
- Exhaustivo (no se puede olvidar ningún posible valor o atributo de la variable).
- Excluyente (nadie puede presentar dos valores o atributos simultáneos de la variable).

1. Naturaleza de los datos

Representan cualidades, atributos o categorías, no cantidades. No se pueden operar matemáticamente (no se pueden sumar los colores o los géneros).

- **Cuantitativa o métrica**

Son aquellas cuya magnitud puede ser medida en términos numéricos, mediante el empleo de un instrumento de medición mecánico (balanza,

flexómetro, termómetro, dinamómetro, pie de rey, peachimetro, Oxigenometro, Multiparamétrico Hanna, etc.).

Pueden ser:

Continua. Sus valores son posibles dentro de cualquier intervalo (infinitos valores fraccionados). Representadas con números enteros (cuando se redondean) o fraccionarios. Se obtienen de mediciones efectuadas con instrumentos mecánicos.

Ejemplos: Altura de la planta (m), hojas activas en una planta de banano

Discreta o discontinua. Puede asumir un número contable de valores. No pueden tomar valores intermedios. No pueden ser fraccionadas. Se obtienen de efectuar un conteo.

Ejemplos: Terneros en el hato ganadero, hojas activas en una planta de banano.

- **Cualitativa o no métrica**

Son atributos que no pueden ser medidos con instrumentos mecánicos. Son medidas con el empleo de instrumentos documentales (solamente cuando se asocian a una frecuencia pueden tratarse de forma estadística).

Pueden ser:

Dicotómica. Son aquellas que tienen dos opciones de respuestas, o sea, dos categorías.

Ejemplos: el sexo de animales, cuyas categorías son machos y hembras. Vivo o muerto. Empleado o desempleado. Iniciador o continuador.

Politómica. Son aquellas que tienen más de dos categorías.

Ejemplos: índice de infestación de una plaga en una granja agrícola determinada, cuyas categorías pueden ser alta, media o baja. Razas de caninos (Chihuahua, Pug, Pastor alemán, Dálmata, etc.). Tipos de sistemas de riego (aspersión, goteo, gravedad, etc.).

2. Escalas de medición

Los datos obtenidos luego de medir una variable presentan distintos atributos de medición dentro de los cuales se encuentran el orden, la distancia y el origen (permiten clasificar a las variables en cuatro escalas de medición).

- **Categórica**

Reciben este nombre porque sus datos son categorías, o sea, atributos.

Pueden ser:

Nominal. No tienen ningún atributo de medición o algún orden en particular. Una categoría no es más que la otra. Se agrupan en categorías mutuamente excluyentes. Las categorías pueden ser nombres como números (funcionan como símbolos).

Ejemplos: sexo de los productores bananeros (femenino, masculino), tipo de alimentación (balanceado, preparación casera), razas de ganado vacuno (Holstein, Brahman, Jersey, Girolando, Brahman).

Ordinal. Son aquellas que cuentan con un orden en sus categorías como único atributo. No se conoce cuanto más es una categoría que la otra. Tienen una jerarquía.

Ejemplos: afectación de una plaga (intensa, media, leve), calidad del servicio en una clínica veterinaria (muy bueno, bueno, regular, inadecuado), etc.

- **Numérica**

Reciben este nombre porque sus datos son unidades, o sea, números.

Pueden ser:

Intervalo. Conserva las características de orden de la escala ordinal y se le agrega el atributo distancia. El 0 es fijado de forma arbitraria (por convención). El 0 es un valor más de la variable.

Puede alcanzar valores negativos.

Ejemplos: Temperatura del agua (0 °C, 20 °C, -5 °C), msnm (5, 500, 1800), número de mascotas de los representantes (0, 1, 2).

Razón. Conserva las características de orden de la escala ordinal y distancia de la escala de intervalo y se le agrega el atributo origen. El 0 es absoluto y representa la ausencia de la variable, pero no del objeto o sujeto. Los valores comienzan a partir del 0.

Ejemplos: Solidos solubles totales del agua, oxígeno disuelto (mg/L).

Las variables numéricas de razón ofrecen mayor cantidad de información debido a que cuentan con los tres atributos de medición, por ello, siempre que sea posible, se debe tratar de obtener la mayor información a partir de recolectarla con la mayor cantidad de atributos, los cuales pueden quitarse si fuese necesario realizar una transformación. Por ello, se puede transformar una escala de medición en otra, siempre que sea en una con menos atributos (escala ordinal en nominal, escala de intervalo o de razón en ordinal o nominal, con el inconveniente de que siempre se presenta pérdida de información (Figura 2.1.3).

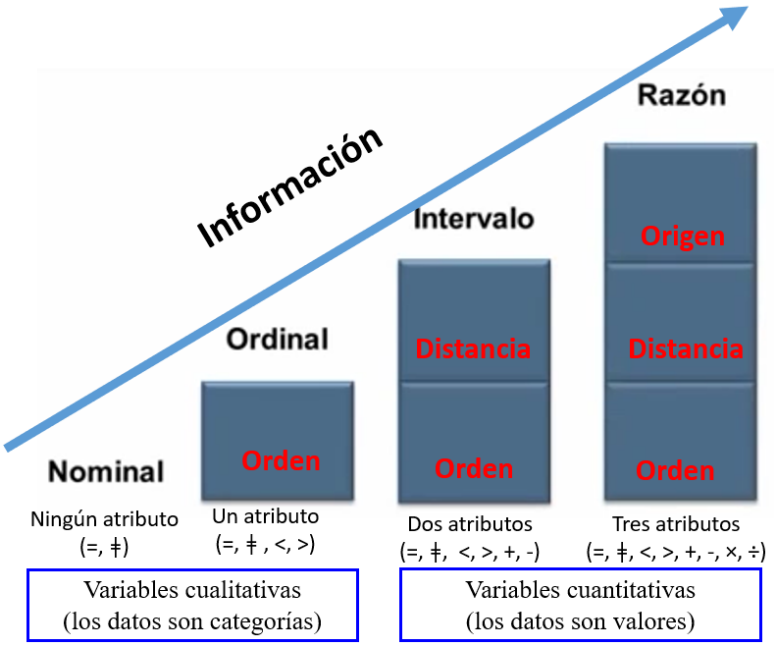


Figura 59. Atributos de medición de las variables categóricas y numéricas.

3. Relación causa-efecto (causalidad)

En un experimento (sea unifactorial o factorial), lo más importante es entender cuál es la causa del cambio y cuál es la respuesta a ese cambio o variabilidad que se manipula intencionalmente. Por eso, las variables se clasifican según su papel dentro de la relación causal entre el tratamiento aplicado y el resultado obtenido (Figura 2.1.4).

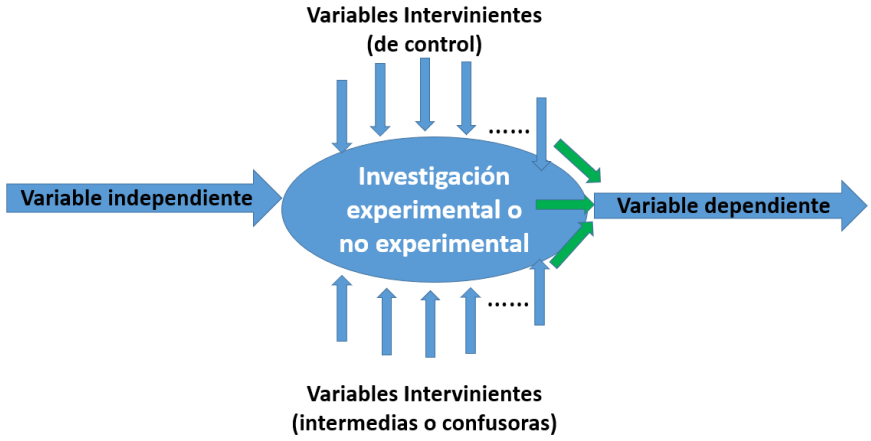


Figura 60. Esquema del proceso experimental con la clasificación de las variables por su relación causa-efecto.

Las variables por la relación causa-efecto pueden ser:

- **Variable independiente**

Es el factor de estudio manipulado por el investigador. Es la causa de variación del fenómeno estudiado (variabilidad intencional y no aleatoria). Pueden manipularse un factor de estudio (experimentos unifactoriales) o dos o más (experimentos factoriales). Es la variable fija ya que el investigador es que realiza la manipulación deliberada y la define previamente. Siempre se ubica en el eje de las abscisas (eje X). Siempre es categórica por su escala de medición. La variable

independiente se encuentra conformada por los niveles o versiones (tratamientos) si es un estudio experimental o los grupos, si es un estudio no experimental.

Ejemplos: Cultivares de maíz, dosis de humus aplicado, tipo de alimento suministrado al ganado, niveles de salinidad en el agua.

- **Variables intervinientes**

Se denominan como variables extrañas. Son no controlados por el investigador, aunque deben ser medidas. Algunas pueden causar confusión. Por su condición se interponen entre la variable independiente y la variable dependiente, y determina sus relaciones. Se ubican en el eje Z.

Pueden ser:

Controladas. Poseen fuerte influencia sobre la variable dependiente, aunque no se estudian como factor causal. Son todos los elementos relacionados con el manejo del experimento y que tienen que mantenerse contantes en todas las unidades experimentales (son fijas) para evitar que afecten a la variable dependiente y generar sesgo, o sea, incrementar la variabilidad. Su control adecuado garantiza que los efectos que se produzcan se deban solamente al FE.

Ejemplos: tipo de suelo, fármacos utilizados, técnicas de manejo agronómico del cultivo, etc.

Confusoras o extrañas. El investigador no las puede controlar de forma adecuada o eliminar y que influyen en la variable dependiente. Afectan la validez interna del experimento. Los resultados pueden mostrar una relación falsa entre causa-efecto ya que aportan variabilidad no controlada. Pueden generar análisis incorrectos de los resultados.

Ejemplos: fertilidad del suelo diferente, procedencia de animales, tipo de especie, etc. Para atenuar su efecto se aplican los principios básicos para una buena experimentación (aleatorización, replicación y técnica de bloqueo).

Intermedias. Se conocen como covariables cuando son numéricas (siempre son aleatorias). Son variables internas del proceso que explican cómo se produce el efecto, aunque no son controladas en el experimento, siempre deben ser medidas.

Ejemplos: Temperatura promedio. Humedad relativa del aire, pH del suelo, ruido, partículas, etc.

En el desarrollo de un experimento la calidad se logra cuando la diferencia que se presente entre los tratamientos se deba a la variable independiente manipulada; además, de controlar las fuentes de variación, mantener las condiciones estables, asegurar independencia entre las unidades experimentales y permitir inferir causa–efecto con validez estadística, todo lo cual, garantiza que los resultados obtenidos sean confiables, repetibles y útiles para la toma de decisiones en ciencias agropecuarias.

- **Variable dependiente**

Es la variable de estudio medida por el investigador con instrumentos mecánicos o documentales. Es la variable respuesta (efecto medido). Siempre se ubica en el eje de ordenadas (eje Y). Su variabilidad se encuentra condicionada por la variable independiente y por las variables intervinientes. Siempre resulta necesario controlar y conocer las variables intervinientes, ya que pueden afectar los resultados y no se puede asegurar que la manipulación de la variable independiente sea lo que ocasiona los cambios en la variable independiente. Si esto ocurre el estudio no es válido.

Ejemplos: Rendimiento agrícola del cultivo de soya (t ha^{-1}), producción de leche por día (L), Vibrios (UFC/mL).

4. Amplitud de las mediciones

Se refiere a la cantidad de sujetos u objetos que se consideran para obtener el dato, o sea, uno o un conjunto (grupo).

Pueden ser

- **Individuales**

Las unidades de observación o de medición son sujetos u objetos. El dato se obtiene de la unidad de estudio de forma directa e individualmente.

Ejemplos: Edad de un productor, altura de planta, peso de camarón, temperatura de un felino, nivel de instrucción de una persona, etc.

- **Colectivas**

Las unidades de observación o de medición son conjuntos o grupos. El dato se obtiene del conjunto de sujetos u objetos que forman parte de la investigación. Ejemplos: PIB de un país, mortalidad infantil, tasa de analfabetismo, densidad poblacional de un cultivo, etc.

Descripción estadística de variables numéricas y categóricas

Medidas de resumen de datos (MRD)

En investigación científica los fenómenos biológicos no suelen ser constantes (en la mayoría de las veces son variables). Por lo tanto, lo primero que se debe conocer en una POE o muestra dada (objetos o sujetos investigados los cuales constituyen las unidades de estudio o de muestreo) son sus parámetros o estadígrafos (estadísticos) correspondientes. Estos elementos permiten realizar (desde el punto de vista estadístico) una descripción tabular a través de la construcción de Tabla de Estadísticos Descriptivos (TED) y Tabla de Distribución de Frecuencias (TDF) o gráfica (cuatro tipos) de una variable numérica o categórica (clasificación por las escalas de medición); a partir del cálculo de las MRD y frecuencias absolutas y relativas; que caracterizan el comportamiento de los datos de una variable dada. Lo anterior me permite realizar una adecuada presentación tabular (tablas) y representación gráfica (gráficos) de los resultados en una investigación científica.

Conocer las diferentes medidas utilizadas para resumir datos numéricos o categóricos, relacionados con sus características, ventajas, desventajas

e importancia práctica constituyen la base teórica y práctica para efectuar una correcta presentación (tablas), representación (gráficos) e interpretación de los resultados obtenidos en una investigación científica (estudio experimental) (Johnson y Kuby, 2012).

Descripción estadística de una variable aleatoria numérica

Medidas de tendencia central

Permiten apreciar el comportamiento medio del conjunto o serie de datos correspondiente a una variable numérica o categórica.

Media aritmética

Es el cociente entre la suma de cada valor de la serie de datos y el número total de elementos que la conforman (se utiliza solamente para variables numéricas). Permite resumir un conjunto o serie de datos en un valor representativo. Se denomina también como promedio. Es la base de las comparaciones múltiples y del análisis descriptivo de los datos.

En una muestra se denota con X barra ($\bar{X}$); y en la población objeto de estudio se denota por la letra griega mu (μ).

Para su cálculo se utilizan las siguientes fórmulas:

Para datos muestrales: $\bar{X} = \frac{X_1+X_2+X_3+\ldots+X_n}{n}$ forma resumida $\bar{X} = \frac{\sum X_i}{n}$

Para datos poblacionales: $\mu = \frac{X_1+X_2+X_3+\ldots+X_N}{N}$ forma resumida $\mu = \frac{\sum X_i}{N}$

Propiedades, ventajas y desventajas

- Es única y para su cálculo se utilizan todos los datos disponibles.
- Representa un valor alrededor del cual oscilan todos los valores de la variable medida (es el punto de equilibrio en un conjunto de datos).

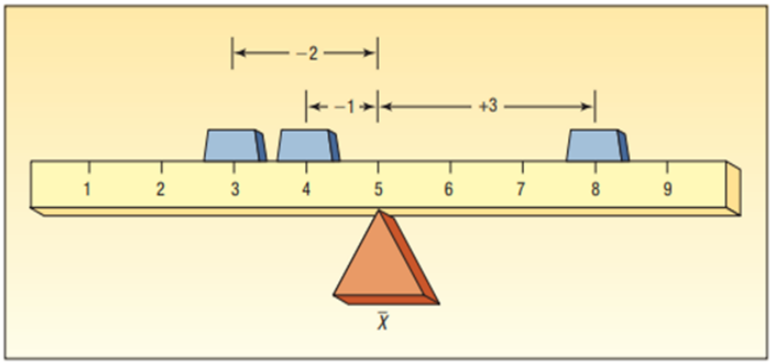


Figura 61.Media aritmética.

- Siempre es utilizada en procedimientos estadísticos para la comparación de medias entre dos o más tratamientos, o entre grupos.
- Su valor puede estar afectado por valores extremos o atípicos (outliers), ya sean altos o bajos (en este caso no es robusta). Es sensible a valores extremos.
- Pierde representatividad en distribuciones de datos asimétricas (media aritmética, mediana y moda diferentes).
- La suma de las desviaciones de cada valor respecto de la media aritmética es cero. La expresión matemática siguiente se utiliza para verificar si la sumatoria de los residuos es igual a 0.

$$\sum = (X_i - \bar{X}) = 0.$$

Mediana (Me)

Supuesta la serie de datos ordenada en forma ascendente o descendente es el valor o categoría que la divide en dos partes iguales (50 % de los datos son inferiores y el otro 50 % son superiores). Es el P_{50}, Q_2 o D_5.

Para su calcula se utilizan dos fórmulas que permiten conocer la posición donde se encuentra el valor de la mediana (si n es impar) o donde se encuentran los dos valores, los cuales se suman y se divide por dos para obtener el valor de la mediada (si n es par). Su uso es más apropiado cuando la variable es categórica ordinal o la numérica presenta valores atípicos.

Fórmulas para determinar la posición de la mediana

Si n es impar $X_{\frac{n+1}{2}}$

Si n es par $X_{\frac{n}{2}}, X_{\frac{n}{2}+1}$

Propiedades, ventajas y desventajas.

- Siempre existe y es única.
- No presenta el problema de estar influenciada por valores extremos, ya que no depende de los valores que toma la variable, sino del orden de los mismas (es apropiado su uso en distribuciones asimétricas).
- Es aplicable a cualquier tipo de datos que puedan ser ordenados, es decir, tiene sentido cuando los datos pueden ser medidos por escalas métricas y ordinales.

- Es probable que la mediana obtenida no sea uno de los valores observados.

Moda (Mo)

Es el valor o atributo de una variable que ocurre o se presenta con la mayor frecuencia, o sea, es el dato que más se repite (frecuencia absoluta).

Propiedades, ventajas y desventajas.

- No se necesita realizar ningún cálculo para obtener el valor o atributo o valores o atributos que más se repiten.
- Se utiliza para variables numéricas y categóricas ordinales.
- Puede no existir un valor modal o atributo (distribución amodal), existir uno (unimodal), dos (bimodal) o más de dos valores modales (multimodal). En este último caso pierde representatividad.
- Para distribuciones simétricas unimodales, la media, la mediana y la moda corresponden al mismo valor.
- Puede estar medianamente afectada por valores extremos; ya que no es afectada por valores extremos individuales, aunque sí puede afectarse si se presentan valores extremos con una frecuencia elevada o si la serie de datos es pequeña,

Medidas de dispersión

Hacen referencia al grado de variabilidad que se presenta en los valores obtenidos en la serie o conjunto de datos.

Rango o recorrido (R o A_T)

Se define como el límite dentro del cual se encuentran comprendidos todos los valores de la serie de datos. Para su cálculo se utiliza la siguiente fórmula:

$$R = V_{máx} - V_{mín}$$

Propiedades, ventajas y desventajas.

- Es intuitivo y útil como descripción rápida de la variabilidad.
- Cuanto mayor es el rango o amplitud de los datos, más dispersos se encuentran alrededor de la media aritmética (es sensible a valores extremos).
- Su valor aumenta con el número de observaciones o se mantiene igual; sin embargo, nunca disminuye (en la recolección de datos).

Varianza

Es el promedio de la sumatoria de las diferencias cuadráticas de n puntuaciones con respecto a su media aritmética.

Para su cálculo se utilizan las siguientes fórmulas:

Para datos muestrales no agrupados: $S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$

Para datos muestrales agrupados: $S^2 = \frac{\sum_{i=1}^{n} f_i(X_i - \bar{X})^2}{n-1}$

Para datos poblacionales no agrupados: $\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$

Para datos poblacionales agrupados: $\sigma^2 = \frac{\sum_{i=1}^{N} f_i(X_i - \mu)^2}{N}$

Propiedades, ventajas y desventajas.

- Expresa el promedio de las desviaciones al cuadrado de las observaciones respecto a la media aritmética.
- Aquellos conjuntos de datos que tengan valores más alejados de la media, sea muestral o poblacional, tendrán una mayor varianza.
- Únicamente adquiere valores mayores o iguales a cero, nunca valores negativos.
- Se usa para la inferencia estadística.
- Presenta el inconveniente que se expresa en la unidad de medida elevada al cuadrado.
- Su valor puede ser elevado y generar confusión en la interpretación de los resultados.

Desviación típica o estándar de los datos

Es el promedio de la desviación de los datos con respecto a la media aritmética.

Se define como la raíz cuadrada (positiva) de la varianza.

Para su cálculo se utilizan las siguientes fórmulas:

Para datos muestrales no agrupados: $S = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$

Para datos muestrales agrupados: $S = \sqrt{\frac{\sum_{i=1}^{n} f_i(X_i - \bar{X})^2}{n-1}}$

Para datos poblacionales no agrupados: $\sigma = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}}$

Para datos poblacionales agrupados: $\sigma = \sqrt{\frac{\sum_{i=1}^{N} f_i(X_i - \mu)^2}{N}}$

Propiedades, ventajas y desventajas.

- Su uso es apropiado para definir la dispersión promedio de los datos alrededor de la media aritmética expresada en las unidades de medida originales.
- Es una medida de variabilidad que toma en cuenta la dispersión real de los datos (valores) respecto a su media aritmética.
- Informa sobre el promedio de las distancias que tienen los datos respecto de su media aritmética.
- Se expresa en las unidades originales de medición de la variable analizada.

Error típico o estándar de la media ($ES_{\bar{X}}$)

Indica en qué medida o cantidad me equivoco en cada DATO obtenido en relación con la media aritmética.

Para su cálculo se utiliza la siguiente fórmula: $ES_{\bar{X}} = \frac{S}{\sqrt{n}}$

Propiedades, ventajas y desventajas.

- Su uso es apropiado cuando se necesita denotar la desviación que se produce en cada dato respecto a su media aritmética.

- Mientras más pequeño sea el error típico o estándar de la media (error aleatorio), más cerca se está de la media aritmética.
- Se utiliza para determinar los intervalos de confianza [Límite inferior (LI) y Límite superior (LS)].

Coeficiente de variación de Pearson (CV)

Es una medida relativa de dispersión que expresa la variabilidad de un conjunto de datos en relación con su media. Es la relación que se presenta entre la desviación típica o estándar de los datos en una muestra o POE y su media aritmética.

Se debe calcular manualmente, ya que Jamovi no tiene implementado esa opción. Para su cálculo se utilizan las siguientes fórmulas:

Para datos muestrales $CV = \frac{S}{\bar{X}} x100$

Para datos poblacionales $CV = \frac{\sigma}{\mu} x100$

Características

- Es adimensional. Las unidades de medida de la desviación estándar de los datos entre la media aritmética se simplifican.
- Se expresa en porcentaje (%), lo que garantiza una mejor interpretación y comunicación de los resultados de las variables medidas. No depende de las unidades de medida.
- Indica la dispersión relativa de los datos alrededor de la media aritmética.

- Si el valor del CV aumenta existe mayor heterogeneidad de los valores de la variable en cuestión; y si disminuye se presenta mayor homogeneidad.

Ventajas

- Facilita comparar la dispersión relativa de los datos entre variables con medias diferentes y medidas en diferentes unidades de medida, o sea, expresadas en diferentes unidades de medición.
- Es fundamental como indicador de la calidad o precisión del experimento (Ruiz-Ramírez, 2009) (Tabla 2.1.3).

Tabla 11. Rangos del CV que establecen la calidad o precisión de un experimento en ciencias agropecuarias.

Rango del CV (%)	Precisión del experimento	Variabilidad de los datos
Menor a 10	Buena	Baja
10-20	Regular	Moderada
20-30	Inadecuada	Alta

Fuente: Pimentel (1990).

Según Patel et al. (2001) si el valor del CV supera el 30 %, los datos deben ser descartados por la baja precisión obtenida en el experimento. No obstante, Gomez y Gomez (1984) y Patel et al. (2001) son más específicos al indicar que los CV varían considerablemente de acuerdo al tipo de experimento, indicando que los rangos aceptables deben ser entre 6 a 8 % para evaluación de cultivares, 10 a 12 % para fertilización y 13 a 15 % para ensayos de evaluación de plaguicidas. Según Gordón-Mendoza y Camargo-Buitrajo (2015) para valorar la precisión de un

experimento se pueden utilizar además del CV, la repetitividad, el error estándar y otros indicadores. Para este autor la repetitividad es el mejor indicador que refleja una buena precisión en el experimento realizado.

Desventajas

- Es sensible a medias aritméticas cercanas a cero. Si la media aritmética es pequeña el CV produce valores distorsionados extremadamente grandes y se vuelve poco confiable.
- Es apropiado, solamente, para variables de razón (donde el cero significa ausencia real), principalmente variables continuas.
- Es menos útil en distribuciones muy asimétricas o con muchos valores atípicos.

El CV de Pearson es una herramienta utilizada para evaluar la dispersión relativa, comparar series o conjuntos de datos distintos y analizar la consistencia de procesos. Su fortaleza principal es su carácter adimensional y porcentual, aunque debe usarse con precaución cuando la media es un valor muy pequeño.

Medidas de posición

Indican el valor por debajo o por encima del cual, se encuentra un determinado porcentaje de datos, dentro del total de observaciones que presenta la serie de datos.

También denominas medidas de orden, ubicación o localización son aquellas que informan sobre distintas características de los datos a partir de la ordenación de los valores observados.

Cuantil estadístico

Son los valores que dividen la serie de datos ordenada de forma ascendente (de menor a mayor) en partes iguales (porcentaje), es decir, en intervalos que comprenden la misma cantidad de valores.

Características

- Para un conjunto de datos el valor de un cuantil dado indica el porcentaje de datos que son iguales, menores o mayores que dicho valor.
- Los cuantiles indican el orden o posición de una observación entre los valores de una variable cuantitativa.
- Para el cálculo de cualquier cuantil se deben ordenar de forma ascendente los valores de la muestra, al resultado de dicha reubicación de los valores se le conoce como muestra ordenada. Mínimo: El mínimo es el valor menor. Máximo: El máximo es el valor mayor.
- Dentro de los cuantiles se encuentran los cuartiles (divide el conjunto de datos en cuatro partes iguales y se forman tres puntos de corte), quintiles (divide el conjunto de datos en cinco partes iguales y se forman cuatro puntos de corte), deciles (divide el conjunto de datos en diez partes iguales y se forman nueve puntos de corte) y percentiles (divide el conjunto de datos en 100 partes iguales y se forman 99 puntos de corte).

Se debe considerar que en una serie de datos los cuantiles pueden ser el mismo valor o no, en dependencia del que se calcula, el Q_1, sería el P_{25}, el D_4 sería el K_2 y el P_{40}, el K_4 sería el D_8 y el P_{80} (Figura 2.1.5).

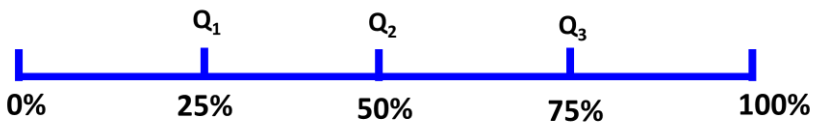


Figura 62. Distribución de los cuartiles en función de los porcentajes de valores que involucran.

Para su cálculo se utilizan las siguientes fórmulas:

Para los cuartiles $Q_i = \frac{\boldsymbol{Q_i(n+1)}}{\boldsymbol{4}}$

Para los quintiles $K_i = \frac{\boldsymbol{K_i(n+1)}}{\boldsymbol{5}}$

Para los deciles $D_i = \frac{\boldsymbol{D_i(n+1)}}{\boldsymbol{10}}$

Para los percentiles $P_i = \frac{\boldsymbol{P_i(n+1)}}{\boldsymbol{100}}$

Para que se utilizan

- Conocer la ubicación de los datos dentro del conjunto.
- Identificar zonas típicas o “normales” en los datos.
- Clasificar valores en niveles (bajo, medio, alto).
- Elaborar el gráfico de caja y sesgos 1-D (Boxplot) (denominado en ocasiones como gráfico de los cinco puntos) y calcular el rango intercuartílico.

Medidas de distribución

Permiten identificar la forma en que se aglomeran o separan los valores de acuerdo a su representación gráfica (distribución teórica de referencia).

Asimetría o sesgo (g_1 o As)

El coeficiente de asimetría de Ficher-Pearson ajustado para el tamaño de la muestra evalúa la forma que adopta la distribución de frecuencias respecto al grado de inclinación que registran los datos respecto al valor promedio tomado como centro de gravedad.

Para su cálculo se utiliza la fórmula:

$$g_1 = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{S}\right)^3$$

Características

- Mide el grado de asimetría de la serie o conjunto de datos con respecto a la media aritmética.
- El coeficiente de asimetría indica si se presenta el mismo número de elementos a izquierda y a la derecha de la media aritmética, o sea, evalúa la proximidad de los datos a su media aritmética.

Tipos de asimetría

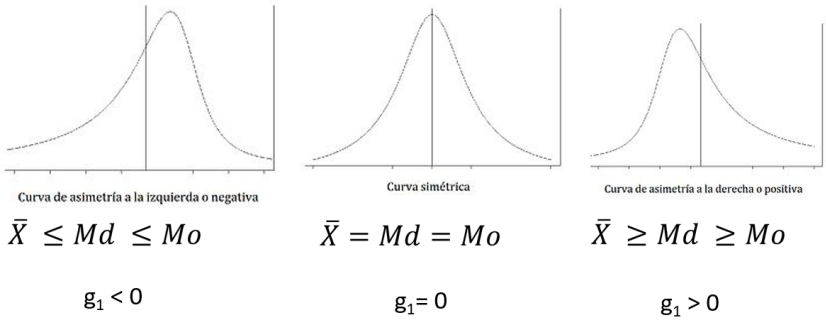


Figura 63. Tipos de asimetría.

Curtosis (g_2)

El coeficiente de curtosis ajustado de Fisher-Pearson mide el grado de agudeza o achatamiento de una distribución de frecuencias con relación a una distribución teórica o de referencia (distribución normal).

Para su cálculo se utiliza la fórmula:

$$g_2 = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{S}\right)^4\right] - \frac{3(n-1)^2}{(n-2)(n-3)}$$

Características

La curtosis (o apuntamiento) es una medida de distribución o forma que mide cuán escarpada o achatada está una curva o distribución.

Tipos de asimetría

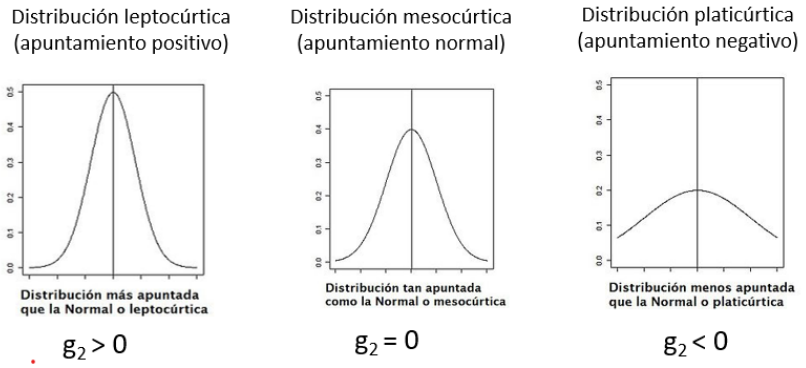


Figura 64. La curtosis.

Frecuencias absolutas y relativas

Indican el número de observaciones [(frecuencias absolutas (conteo o recuento) y relativas (%)] dentro de cada categoría que puede alcanzar la variable. Es posible convertir las frecuencias de las categorías o de la clase (absolutas), en frecuencias relativas de clase, para mostrar la fracción del número total de observaciones en cada una de ellas. Así, una frecuencia relativa capta la relación entre la totalidad de elementos de una categoría o clase y el número total de observaciones.

Se entiende por frecuencia absoluta a la cantidad de veces que una categoría o un valor de una variable es medido en un conjunto de datos y distribución de frecuencias es el método estadístico que se utiliza para describir dicho conjunto de datos.

Tipos de frecuencias que conforman la TDF

1. Frecuencia absoluta (n_i). Es el número de repeticiones que presenta una observación (cantidad o recuento).
2. Frecuencia relativa (f_i). Es la frecuencia absoluta dividida por el número total de datos (porcentaje).
3. Frecuencia absoluta acumulada (N_i). Es la suma de los distintos valores de la frecuencia absoluta tomando como referencia un punto dado (la última frecuencia absoluta acumulada es igual al número de casos).
4. Frecuencia relativa acumulada (F_i). Es el resultado de dividir cada frecuencia absoluta acumulada por el número total de datos.

Características de la TDF

- Es una tabla que divide un conjunto de datos en un número adecuado de "clases" o categorías.
- Se utilizan para presentar los datos obtenidos en alguna investigación en forma clara y ordenada.
- Son autoexplicativas.
- La TDF para variables numéricas carece de importancia práctica.

La ventaja principal de la TDF es que se utilizan en datos de cualquier escala de medición, aunque, es más apropiada para una variable categórica.

Representación gráfica de resultados obtenidos en variables numéricas o categóricas

Las medidas de resumen de datos ofrecen una primera aproximación numérica a los datos y permiten resumir sus rasgos más importantes; sin

embargo, cuando el objetivo es comprender los datos de manera más intuitiva o comunicar hallazgos de forma clara y accesible, los recursos gráficos cobran un papel decisivo. Un gráfico puede revelar patrones que pasan inadvertidos en las tablas, mostrar tendencias emergentes o incluso advertir sobre valores atípicos que podrían influir en los análisis posteriores. Los gráficos permiten visualizar lo que los números apenas insinúan, facilitando la exploración inicial del conjunto de datos y brindando una comprensión más completa del fenómeno estudiado. Por eso, en cualquier proceso de análisis y comunicación científica las representaciones visuales se convierten en herramientas fundamentales para conectar los resultados con el lector y guiarlo hacia una interpretación más profunda y cercana de la información.

La representación gráfica de los resultados de una variable aleatoria numérica o categórica se realiza de forma diferente. Pueden construirse gráficos univariados (Tabla 2.1.4), bivariados y multivariados en función de la cantidad de variables que se consideren.

Tabla 12. Resumen de los principales gráficos univariados que se realizan por tipo de variable.

Tipo de variable		**Tipo de gráfico recomendado**	**Para que se utiliza**
Escala de medición	**Naturaleza de los datos**		
Numérica (intervalo o razón)	Cuantitativa discreta	Histograma simple	Ver distribución
	Cuantitativa continua	Diagrama de caja y sesgos 1-D	Ver variabilidad

Categórica (nominal u ordinal)	Cualitativa (politómica (cuatro o más categorías) Cualitativa (dicotómica (dos categorías) y politómica (tres categorías)	Gráfico de barras simples Gráfico circular	Frecuencias absolutas y relativas Proporciones

Análisis. Al hacer clic en el menú análisis del Jamovi se puede acceder a la opción exploración-descriptivas donde se despliega la viñeta de gráficos y se muestran varias alternativas de gráficos univariados en función del tipo de variable. Histograma simple, diagrama de caja y sesgos, gráfico de barras simples y gráfico circular.



Figura 65.Análisis.

Gráficos para variables numéricas

Histograma simple

Sobre el eje X se colocan los distintos intervalos o clases y sobre cada uno de ellos se levanta un rectángulo de altura igual a la frecuencia observada (absoluta o relativa) del intervalo.

Características

Es equivalente al diagrama de barras simples, aunque se utiliza para datos numéricos, de tipo continuo (peso o talla) o discreto (edad en años), si tiene un gran número de datos y los valores son continuos se categorizan en clases o intervalos (Figura 2.1.6A). Cuenta con la opción de colocarle la curva de la distribución para conocer el grado de variabilidad (Figura 66).

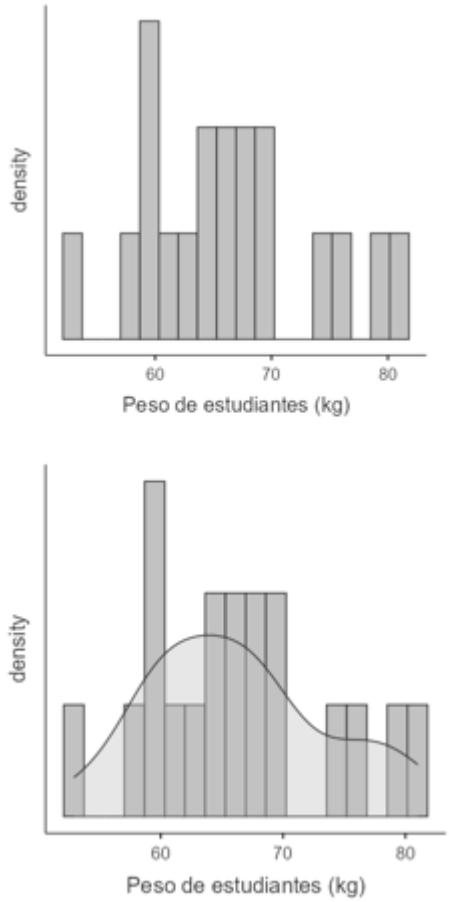


Figura 66.A. Histograma simple de la variable peso de estudiantes. B. Histograma con la curva de distribución de los datos.

Diagrama de cajas

Es una representación gráfica de los datos que permite analizar conjuntamente una serie de medidas numéricas como: límite inferior, límite superior y los cuartiles (Q_1 , Q_2 que representa a la mediana y Q_3), por ello, se le conoce como gráfico de los cinco puntos. Denominado de forma técnica como Diagrama de caja y sesgos 1D.

Características

Se utiliza cuando se estudia una variable numérica o grupos diferentes evaluados con una variable numérica.

El diagrama de caja y sesgos (Figura 2.1.7) facilita:

- Observar características de la distribución de datos (simetría).
- La presencia de valores atípicos que afectan la distribución.
- Comparar la dispersión de dos o más conjuntos de datos.

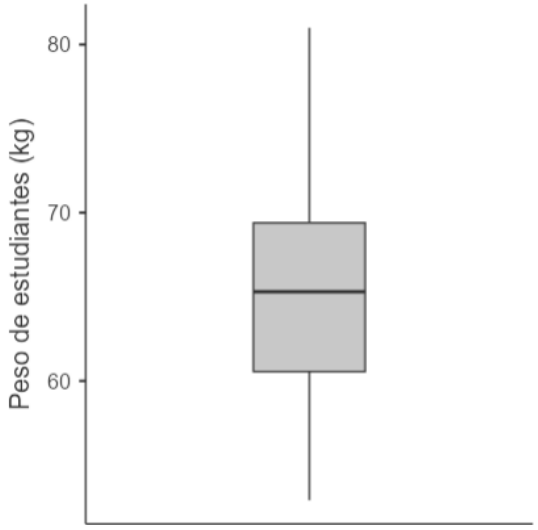


Figura 67.Diagrama de caja y sesgos 1-D de la variable peso de estudiantes.

Gráficos en variables categóricas

Representa las frecuencias (absolutas o relativas) de las categorías o intervalos sin agrupar dispuestas de forma paralela, de forma que la

extensión de cada barra es proporcional a la magnitud que presenta dicha categoría.

Características

El gráfico de barras simples (Figura 2.1.8) se utiliza para representar variables categóricas (nominal u ordinal) o numéricas (cuando los datos se agrupan en intervalos o clases) politómicas con cuatro o más categorías.

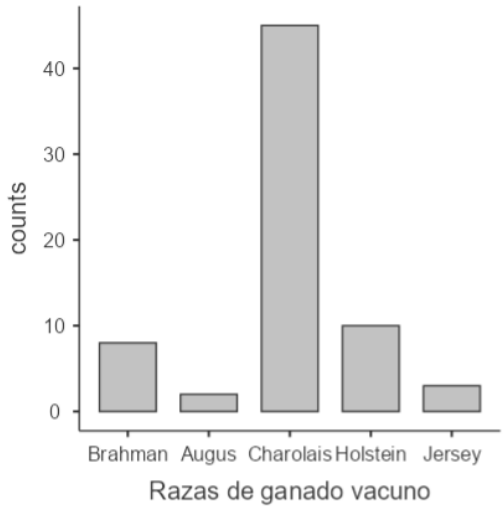


Figura 68. Gráfico de barras simples de la variable razas de ganado vacuno.

Diagrama circular o de sectores

Consiste en un círculo dividido en sectores circulares o semicirculares que se corresponden con las distintas categorías o intervalos de la variable. El área de cada sector es proporcional a la frecuencia absoluta o relativa de cada categoría o clase.

Características

Se utiliza para representar cualquier tipo de variable (nominal, ordinal o numérica con intervalos o clases (politómicas) que tengan cuatro o menos de cuatro categorías (Figura 2.1.9).

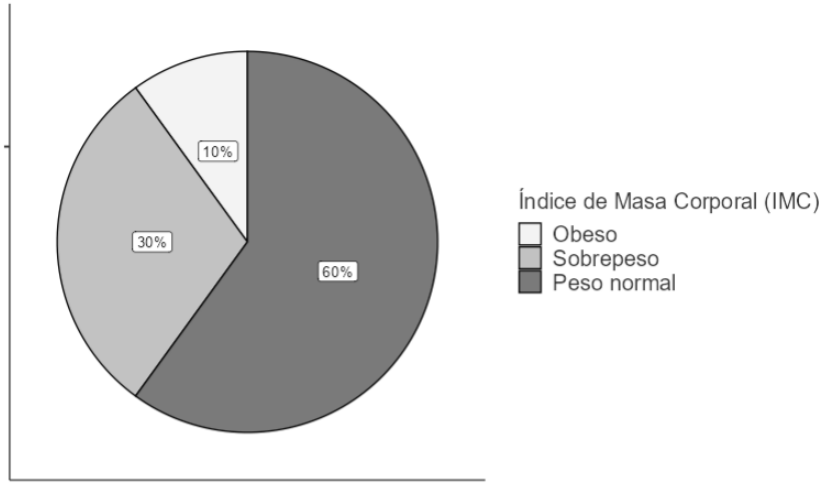


Figura 69. Gráfico circular de la variable IMC.

2.1.5 Otras opciones para la construcción de gráficos univariados con Jamovi

Jamovi cuenta con opciones adicionales mediante la instalación de módulos 

al lado derecho superior, en la barra de herramientas, de la interfaz del Software.

El módulo JJStatsPlot 

una vez instalado y visible en la barra de herramientas, ofrece varias opciones de *gráficos univariados*:

1. Para variable numérica continua: Histograma simple.

2. Para variable categórica nominal u ordinal (Gráfico circular).

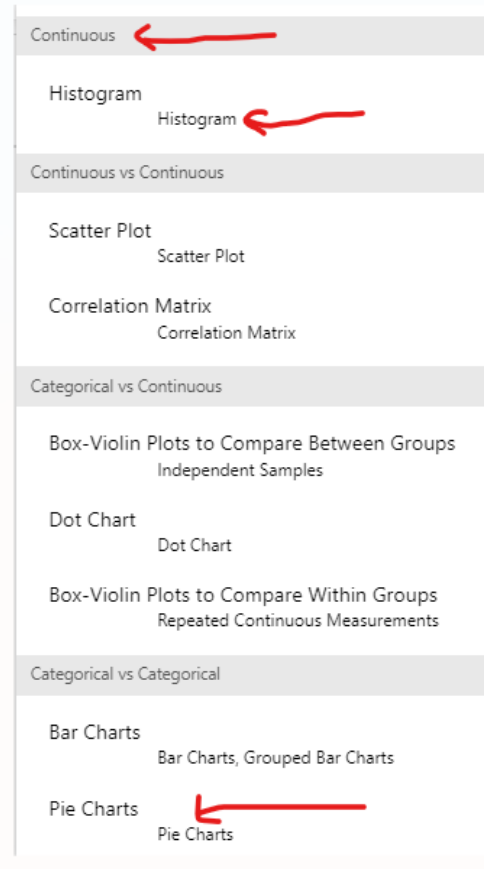


Figura 70. Módulo JJStatsPlot.

Conclusiones

La historia de la Estadística refleja la evolución de la humanidad en su necesidad de contar, organizar y procesar datos. Desde los censos antiguos hasta los algoritmos modernos, la Estadística ha pasado de ser un instrumento administrativo a convertirse en una ciencia aplicada de alcance universal, necesario para enfrentar los desafíos del mundo contemporáneo.

Reconocer y dominar los conceptos básicos de Estadística es esencial debido a que permiten interpretar resultados correctamente, tomar decisiones fundamentadas y evitar errores en investigaciones científicas. Sin ellos, la información puede ser incorrectamente utilizada y conducir a conclusiones equivocadas.

Las variables clasificadas por su escala de medición, aumentan el nivel de información desde las categóricas nominales hasta las numéricas continuas; debido a que estas últimas presentan mayor cantidad de atributos de medición. La selección, definición y empleo de variables resultan fundamentales, porque determinan la capacidad del estudio para responder a la hipótesis planteada y los objetivos planificados; además, condicionan el análisis estadístico de los datos que se realizará.

La descripción de variables constituye una de las principales funciones de la estadística, la misma puede ser realizada por medio del cálculo de las medidas de resumen de datos, en los cuales se muestra la forma en que se comporta o descubrir patrones de distribución ocultos en la información recolectada previamente.

La TDF es una herramienta que ayuda a una mejor comprensión de los datos obtenidos en escalas nominal y ordinal.

La representación gráfica de datos de una variable categórica se efectúa de forma más adecuada con un gráfico de barras simples o un gráfico circular; sin embargo, para las variables numéricas la opción más adecuada es la utilización de un histograma simple o un diagrama de caja y sesgos.

La elección de la opción para presentar o representar datos categóricos o numéricos, ya sea de forma tabular o gráfica, depende del criterio del investigador, en función de la búsqueda del mejor entendimiento por parte del público objetivo.

2.2 Práctica 3. Descripción tabular y representación gráfica de variables numéricas y categóricas

Contenidos

Matrices de datos. Tabla de estadísticos descriptivos. Tabla de distribución de frecuencias. Representación gráfica de resultados de variables numéricas y categóricas. Empleo de software estadístico Jamovi.

Objetivos

1. Describir de forma tabular y representar gráficamente un conjunto de datos procedentes de variables numéricas y categóricas con la utilización de las opciones que facilita el software estadístico Jamovi.

Introducción

La descripción estadística de variables constituye una de las principales funciones de la Estadística, la misma puede ser realizada por medio del cálculo de las medidas de resumen de datos, tablas y gráficos, en los cuales se muestra la forma en que se comporta o descubrir patrones de distribución ocultos en la información recolectada previamente.

Sin embargo, no todos los procedimientos estadísticos son realmente útiles para las diferentes escalas de medida: lo que se debe tener en cuenta en el momento de realizar un análisis descriptivo.

El software estadístico puede adquirir datos de cualquier tipo de archivo y utilizarlos para generar informes tabulares, construcción de gráficos y diagramas de distribución. Las representaciones gráficas se convierten en un medio eficaz para la interpretación de los resultados, ya que las regularidades se recuerdan con más facilidad cuando se observan gráficamente.

2.2.1 Ejercicio 3.1 Creación de matrices de datos con Jamovi.

Contexto

Se dispone diferentes tipos de variables relacionadas con las Ciencias Agropecuarias, las cuales se muestran a continuación.

1. Fertilizantes orgánicos (humus de lombriz, compost, lixiviado).
2. Temperatura del agua (°C).
3. Altura de la planta de alfalfa (cm).
4. Días a la floración del cultivo de maní.
5. Estado civil de los productores de banano (soltero, casado, divorciado, unión libre, viudo, etc.).
6. Índice de infestación por Vibrios (alto, medio, bajo).
7. Largo de la vaina de haba (cm).
8. Rendimiento agrícola del maíz (t ha^{-1}).
9. Tipo de cosecha empleada (manual, semimecanizada, mecanizada).

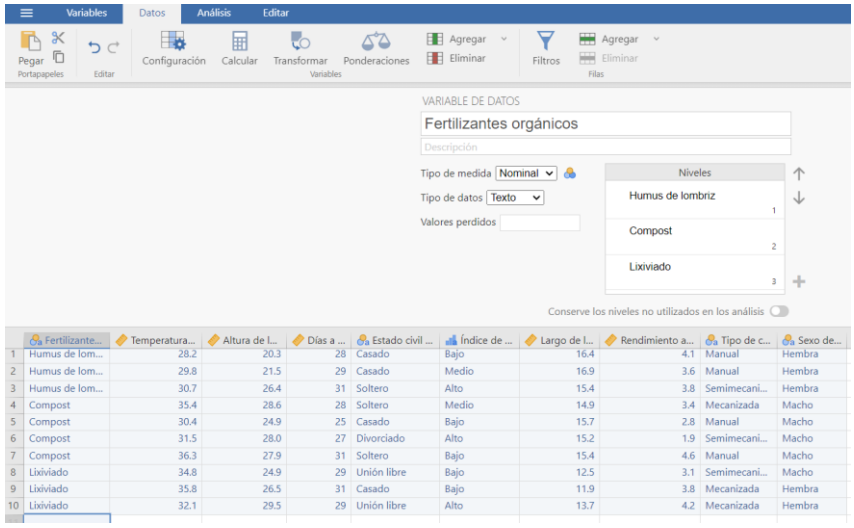
10. Sexo de la mascota (hembra o macho).

Requerimiento del ejercicio

1. Ordene y construya la matriz de datos donde se ubiquen las 10 variables indicadas anteriormente. Por cada una de ellas coloquen 10 datos (creados por usted).

Procedimiento

1. Abrir el software, renombrar el archivo (ejercicio 3.1) y ubicarlo en la carpeta creada en su ordenador para ubicar los ejercicios.
2. Hacer clic en la barra de menú opción editar para colocar el nombre, las características de las variables y guardar lo realizado.
3. Hacer clic en el menú datos; para cargar los 10 datos de cada variable.



	Fertilizante...	Temperatura...	Altura de l...	Días a ...	Estado civil ...	Índice de ...	Largo de l...	Rendimiento a...	Tipo de c...	Sexo de...
1	Humus de lom...	28.2	20.3	28	Casado	Bajo	16.4	4.1	Manual	Hembra
2	Humus de lom...	29.8	21.5	29	Casado	Medio	16.9	3.6	Manual	Hembra
3	Humus de lom...	30.7	26.4	31	Soltero	Alto	15.4	3.8	Semimecani...	Hembra
4	Compost	35.4	28.6	28	Soltero	Medio	14.9	3.4	Mecanizada	Macho
5	Compost	30.4	24.9	25	Casado	Bajo	15.7	2.8	Manual	Macho
6	Compost	31.5	28.0	27	Divorciado	Alto	15.2	1.9	Semimecani...	Macho
7	Compost	36.3	27.9	31	Soltero	Bajo	15.4	4.6	Manual	Macho
8	Lixiviado	34.8	24.9	29	Unión libre	Bajo	12.5	3.1	Semimecani...	Macho
9	Lixiviado	35.8	26.5	31	Casado	Bajo	11.9	3.8	Mecanizada	Hembra
10	Lixiviado	32.1	29.5	29	Unión libre	Alto	13.7	4.2	Mecanizada	Hembra

Figura 71.Estadísticos descriptivos en variables numéricas.

Contexto

Se dispone de los datos muestrales del peso (kg) de 30 caninos atendidos durante una semana en la Clínica de Especialidades Veterinarias de la Universidad Técnica de Machala.

Datos recolectados

Tabla 13. Datos recolectados.

No. de mascota	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Peso (kg)	22,5	12,3	3,5	4,8	35,7	3,9	15,6	7,8	23,6	15,7
No. de mascota	11	12	13	14	15	16	17	18	19	20
Peso (kg)	26,9	4,2	22,6	25,1	15,3	4,6	14,8	27,3	32,1	14,9
No. de mascota	21	22	23	24	25	26	27	28	29	30
Peso (kg)	3,7	61,0	34,9	4,9	25,4	23,2	27,3	22,9	4,9	13,8

Requerimientos del ejercicio

- Construya la matriz de datos correspondiente, renombre el archivo y guárdelo en su computador.
- Realice una caracterización estadística del comportamiento del peso de caninos mediante el empleo de las medidas de resumen de datos.
- Construya un histograma con la curva de distribución de datos.

- Clasifique la serie de datos en relación a la asimetría y curtosis.

Procedimiento

1. Abrir el software, renombrar el archivo (ejercicio 3.2) y ubicarlo en la carpeta creada en su ordenador para ubicar los ejercicios, editar la variable con sus características e ingresar los datos.
2. Hacer clic en la barra de menú opción análisis, exploración, descriptivas, pasar la variable peso de caninos hacia el recuadro variables al lado derecho, dejar variables en columnas para que la tabla de estadísticos descriptivos (TED) se construya de forma vertical y marcar todas las opciones correspondientes a las medidas de resumen de datos.



Figura 72. Software estadístico.

El software estadístico genera una tabla de estadísticos descriptivos que puede copiarse en forma de imagen o de forma editable en una hoja de Word. En la edición de la tabla se deben considerar ubicar la n minúscula que representa que se trabaja con una muestra, colocar la coma en los decimales, ajustar el valor de la moda ya que el programa emite el valor más pequeño cuando se presentan conjuntos de datos con dos o más modas (Tabla 14).

Tabla 14. Tabla de estadísticos descriptivos de la variable peso de caninos.

Estadísticos o estadígrafos	Peso de caninos (kg)
n	30
Perdidos	0
Media aritmética	18,50
Error típico o estándar de la media	2,35
IC 95 % de la media límite inferior	13,69
IC 95 % de la media límite superior	23,31
Mediana	15,65
Moda	4,90 y 27,3
Desviación típica o estándar de los datos	12,88
Varianza	165,98
RIC	19,70
Recorrido o Rango	57,50
Mínimo	3,50
Máximo	61,00
Asimetría	1,16
Curtosis	2,55
25percentil (Q_1)	5,63
50percentil (Q_2)	15,65
75percentil (Q_3)	25,33

Nota. El Intervalo de Confianza de la media supone que las medias muestrales siguen una distribución t con N - 1 grados de libertad.

3. Para construir el gráfico indicado hacer clic en gráficos y marcar histograma y densidad para mostrar la curva de la distribución de datos.

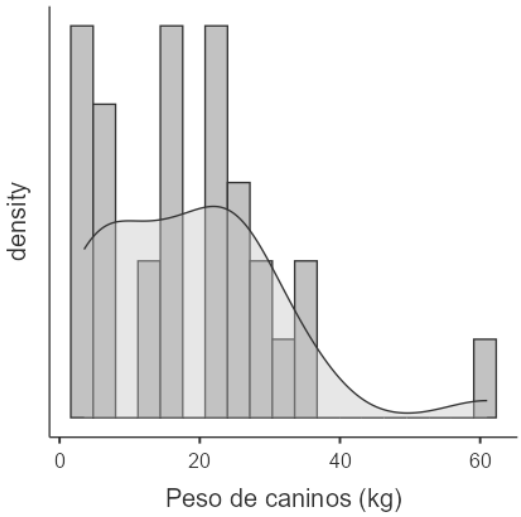


Figura 73. Asimetría.

4. La serie de datos por la asimetría se clasifica como asimétrica positiva debido a que la g_1=1,16 kg y por la curtosis es leptocúrtica ya que la g_2= 2,55 kg.

2.2.2 Ejercicio 5 Estadísticos descriptivos en variables categóricas

Contexto

Se dispone de los datos muestrales del peso de caninos categorizadas (bajo peso, peso normal y sobrepeso) de 30 caninos atendidos durante una semana en la Clínica de Especialidades Veterinarias de la Universidad Técnica de Machala.

Datos recolectados

Tabla 15. Estadísticos descriptivos en variables categóricas.

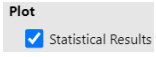
No. de mascota	Categorías	No. de mascota	Categorías	No. de mascota	Categorías
1	Normal	11	Normal	21	Bajo peso
2	Bajo peso	12	Bajo peso	22	Sobrepeso
3	Bajo peso	13	Normal	23	Sobrepeso
4	Bajo peso	14	Normal	24	Bajo peso
5	Sobrepeso	15	Bajo peso	25	Normal
6	Bajo peso	16	Bajo peso	26	Normal
7	Bajo peso	17	Bajo peso	27	Sobrepeso
8	Bajo peso	18	Normal	28	Normal
9	Normal	19	Sobrepeso	29	Bajo peso
10	Bajo peso	20	Bajo peso	30	Bajo peso

Requerimientos del ejercicio

1. Construya la matriz de datos correspondiente, renombre el archivo y guárdelo en su computador.

2. Realice una caracterización estadística del peso de caninos categorizado mediante las frecuencias absolutas y relativas (construir TDF).
3. Construya un gráfico circular.

Procedimiento

1. Abrir el software, renombrar el archivo (ejercicio 3.3), ubicarlo en la carpeta creada en su ordenador, editar la variable con sus características e ingresar los datos.
2. Hacer clic en la barra de menú opción análisis, exploración, descriptivas, pasar la variable peso de caninos (categorizada) hacia el recuadro variables que está al lado derecho, marcar el recuadro de Tabla de frecuencias (TDF); y en el visor de resultados del Jamovi se presenta la TDF, la cual debe ser editada para efectuar la publicación de un artículo científico ya que no puede ser editada.
3. Hacer clic en la barra de menú opción análisis, módulo JJStatsPlot. Marcar en Pie Charts, pasar la variable peso de caninos (categorizada) hacia el recuadro de variable dependiente, al lado derecho, marcar Plot y el recuadro de resultados estadísticos 

y en el visor de resultados del Jamovi se presenta el gráfico circular.

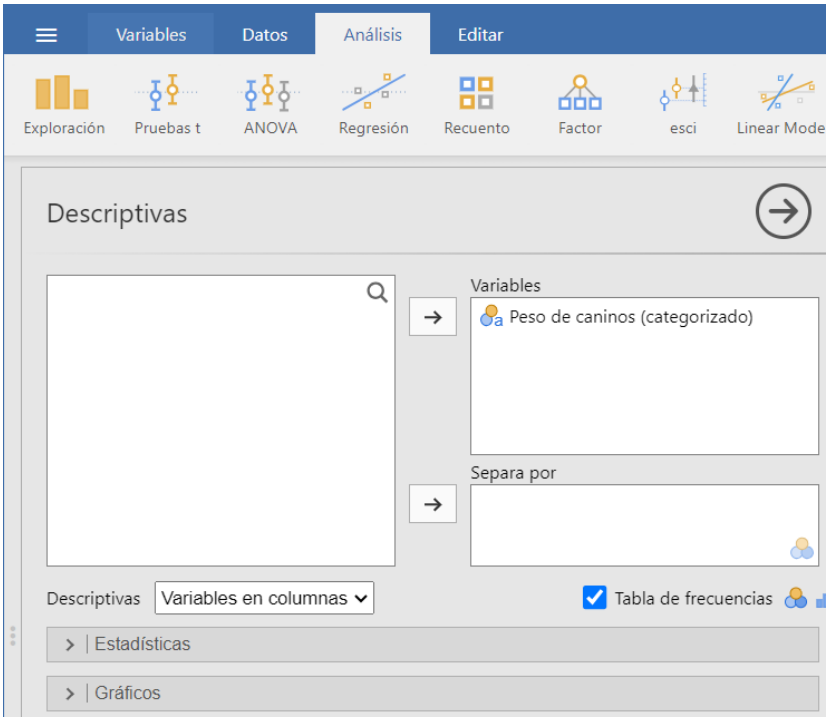


Figura 74.Gráfico circular.

Frequencias de Peso de caninos (categorizado)

Peso de caninos (categorizado)	Frecuencias	% del Total	% Acumulado
Normal	9	30.000 %	30.000 %
Bajo peso	16	53.333 %	83.333 %
Sobrepeso	5	16.667 %	100.000 %

Figura 75.Frecuencia relativa.

Con fines de publicación en artículos científicos o trabajos de titulación la TDF que se genera en Jamovi debe ser editada, debiéndose considerar que las categorías se ubican en el orden de mayor a menor en función de la frecuencia relativa (Tabla 16).

Tabla 16. Distribución de frecuencias de la variable peso de caninos categorizada.

Categorías	Frecuencia absoluta	Frecuencia relativa (%)	Frecuencia absoluta acumulada	Frecuencia relativa acumulada (%)
Bajo peso	16	53,3	16	53,3
Normal	9	30,0	25	83,3
Sobrepeso	5	16,7	30	100
Total	**30**	**100**		

Otra opción es representar los resultados mediante un gráfico circular, considere que si se presentan en la TDF no deben representarse en un gráfico y viceversa (Figura 76).

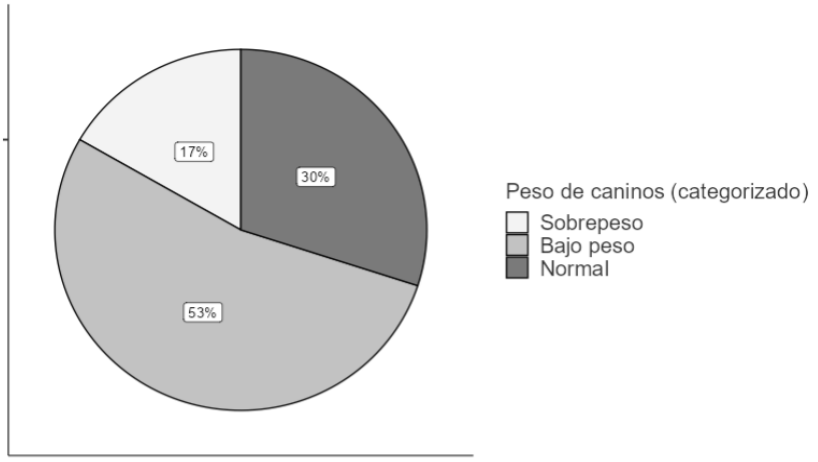


Figura 76. Representación gráfica de los resultados obtenidos en la variable peso de caninos (categorizada).

Conclusiones

La descripción de datos numéricos se realiza a través del cálculo de las medidas de resumen de datos y constituye la base de la Estadística univariada.

La representación gráfica de datos de variables categóricas se efectúa de forma más adecuada con diagramas de barras simples y circular, sin embargo, para las variables numéricas la opción más adecuada es la utilización de histograma simple y diagramas de caja y sesgos.

La elección de la opción para presentar o representar datos categóricos o numéricos, ya sea de forma tabular o gráfica, depende del criterio del investigador, en función de la búsqueda del mejor entendimiento por parte del público objetivo.

2.3 Estadística inferencial

2.3.1 Contenidos

Inferencia estadística. Hipótesis. Características. Prueba de hipótesis. Procedimiento de prueba de hipótesis. Supuestos del modelo paramétrico. Pruebas paramétricas. Test t. Prueba t de Student para grupos independientes. Test F. Análisis de varianza de un factor intertratamientos. Análisis de varianza de un factor intratratamientos. Análisis de varianza factorial intertratamientos. Pruebas Post hoc. Representación gráfica de resultados.

2.3.2 Objetivos

1. Comprender los principios fundamentales que caracterizan el procedimiento de prueba de hipótesis paramétricas que posibilite la obtención de un juicio de valor y la toma de decisiones con fundamento matemático en contextos científicos y productivos.
2. Aplicar de manera sólida los fundamentos teóricos y prácticos de la prueba t de Student para tratamientos independientes, análisis de varianza de un factor intertratamientos, análisis de varianza de un factor intratratamientos y del análisis de varianza factorial intertratamientos, reconociendo su estructura, supuestos y cuando utilizarlas.

2.3.3 Introducción

La Estadística descriptiva caracteriza las variables en un estudio mediante las medidas de resumen de datos; sin embargo, la Estadística inferencial busca tomar decisiones sobre la POE en base a los datos obtenidos en muestras. Una prueba de hipótesis contrasta una afirmación sobre un parámetro (media, varianza, proporción, etc.), cuantificando la evidencia contra H_0 mediante un estadístico de prueba y un p-valor.

En experimentación agropecuaria cuando se manipula un factor de estudio con dos o más tratamientos y se necesita comprobar estadísticamente (con un fundamento matemático) si se presentan o no diferencias estadísticas significativas entre cada tratamiento (se denominan grupos en estudios no experimentales), en función de la variable dependiente medida en las unidades de estudio o de muestreo)

se desarrolla un contraste de prueba de hipótesis; el cual se desarrolla mediante un procedimiento denominado *Ritual de la significancia estadística.* Siempre que se necesite obtener un juicio de valor se utiliza este procedimiento, el cual es metodológico (método científico), estadístico (herramienta) y hermenéutico (interpretativo).

En una investigación científica se necesitan realizar comparaciones, probar ciertas ideas que se tienen alrededor del comportamiento de uno, dos o más tratamientos o grupos, para ello, se necesita realizar el contraste de prueba de hipótesis estadística.

Intención del estudio

En investigación científica es importante conocer que no todos los estudios se realizan con el desarrollo de una prueba de hipótesis, ya que depende de la intención analítica y se relaciona con el enunciado del estudio que defina el investigador (Figura 2.2.1). Considerando que en un estudio experimental lo más común es comprobar las ideas de investigación que se tienen, se debe aplicar el procedimiento de prueba de hipótesis, por ello, en relación con esta temática se profundizará en el presente libro.

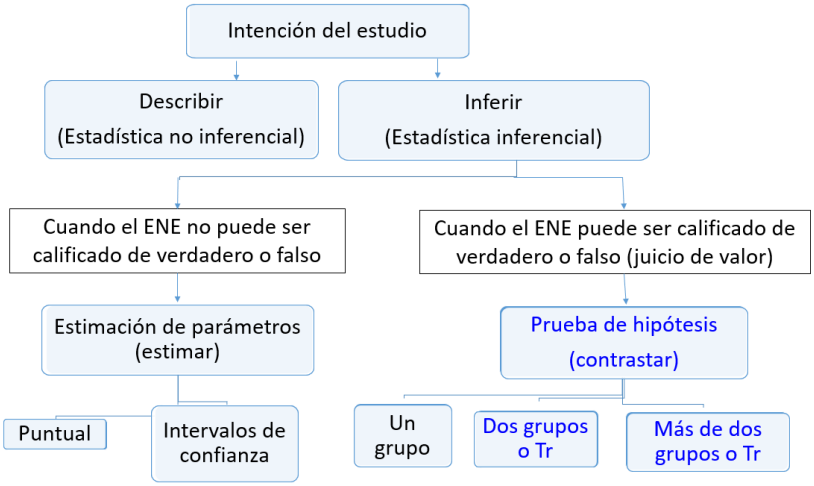


Figura 77. Flujograma que explica el camino a seguir en función de la intención analítica del investigador y el procedimiento estadístico.

Hipótesis. Definición (desde el punto de vista de la metodología de la investigación)

Es una afirmación (suposición, proposición, conjetura, asunción) anticipada del investigador que busca explicar un fenómeno, su validez debe comprobarse mediante la investigación o el razonamiento lógico. Puede ser calificada de verdadera o falsa (se emite un juicio de valor). En el ámbito científico, constituye el punto de partida para formular preguntas y diseñar experimentos. Por su carácter provisional (no es una verdad absoluta), debe ser verificada o refutada mediante pruebas empíricas.

Características de una hipótesis

- **Claridad y precisión:** formulación comprensible y específica.

- **Comprobabilidad:** someterse a verificación mediante observación o experimentación.
- **Fundamentación:** se apoya en conocimientos previos, teorías o datos existentes.
- **Carácter predictivo:** anticipa los resultados o comportamientos de las variables estudiadas.

La hipótesis es el núcleo del método científico, transforma una suposición en una afirmación que es sometida a una prueba. Garantiza la dirección y el propósito de la investigación.

Prueba de hipótesis. Concepto (desde el punto de vista de la Estadística)

Es un procedimiento metodológico-estadístico-interpretativo utilizado para verificar o contrastar la veracidad (verdadera) o falsedad (falsa) de una hipótesis, en base a los datos obtenidos en las unidades de estudio o de muestreo procedentes de una muestra aleatoria (evidencia muestral y teoría de la probabilidad de ocurrencia de un fenómeno o evento) procedente de la POE correspondiente.

Cuestiones fundamentales para desarrollar un procedimiento de prueba de hipótesis

1. Definir sobre que parámetros (POE) se plantean las hipótesis [hipótesis nula (H_0) y la hipótesis alterna (H_1)], es decir, si la prueba de hipótesis es sobre la media [(μ), o la varianza (σ^2) (se utilizan para variable numérica y se aplican pruebas

paramétricas)], o sobre la proporción [(*P*) (se utiliza para variable categórica y se aplican pruebas no paramétricas)] poblacionales.

2. Establecer si es una prueba de hipótesis en un tratamiento o grupo con una medición, un tratamiento o grupo con dos mediciones, en dos tratamientos o grupos, en más de dos tratamientos o grupos independientes y en tres o más tratamientos o grupos medidos en diferentes momentos.

En un tratamiento o grupo: cuando se compara la μ, σ^2, o *P* de un tratamiento o grupo, con un valor de prueba especificado, definido previa investigación teórica).

En un tratamiento o grupo medido dos veces: cuando se compara la μ, σ^2, o *P* de un tratamiento o grupo antes de la intervención o acontecimiento; con la μ, σ^2, o *P* del otro tratamiento o grupo después de la intervención o acontecimiento.

En dos tratamientos o grupos: cuando se compara la μ, σ^2, o *P* de un tratamiento o grupo con la μ, σ^2, o *P* del otro tratamiento o grupo.

En más de dos tratamientos o grupos: cuando se compara la μ, σ^2, o *P* de cada tratamiento o grupo con la μ, σ^2, o *P* general (gran media).

En más de dos tratamientos o grupos y se efectúan las mediciones en las mismas unidades: cuando se compara la μ, σ^2, o *P* de cada tratamiento o grupo con la μ, σ^2, o *P* general en cada momento (antes, durante, después).

Cuando se manipulan dos o más factores de estudio con dos o más tratamientos cada uno: se comparan la μ, σ^2, o *P* de cada combinación

de tratamientos con la μ, σ^2, o *P* general para determinar si se presenta o no efecto de interacción entre los tratamientos de cada factor de estudio.

Procedimiento para el desarrollo de una prueba de hipótesis

El procedimiento de prueba de hipótesis constituye uno de los pilares fundamentales de la inferencia estadística, que permite tomar decisiones objetivas sobre parámetros poblacionales a partir de información muestral. A través de este proceso, el investigador plantea inicialmente una hipótesis nula, que representa una afirmación de ausencia de efecto o de igualdad, y una hipótesis alterna, que expresa la posibilidad de un cambio, diferencia o relación significativa. El método estadístico evalúa la evidencia que aportan los datos frente a estas afirmaciones. Se considera la variabilidad inherente al muestreo y se cuantifica el grado de incertidumbre mediante estadísticos de prueba y p-valores.

El procedimiento de prueba de hipótesis ofrece un marco riguroso y sistemático para respaldar decisiones científicas, mediante el cual se evita emitir conclusiones basadas únicamente en apreciaciones subjetivas y permite interpretar los resultados en función del contexto y los objetivos del estudio planificados.

Los tres primeros pasos corresponden a la metodología de la investigación, los tres siguientes a la Estadística (se debe considerar que el saber hacer o know-how al desarrollar la ejecución de la prueba con el uso de Jamovi no se explica en el proyecto, articulo o trabajo de titulación, ya que se relaciona con las habilidades y experiencia práctica del investigador) y finalmente la hermenéutica, o sea, a la interpretación

del fenómeno estudiado, con un argumento matemático-científico (Figura 78).



Figura 78. Flujograma que explica el camino a seguir en función de la intención analítica del investigador y el procedimiento estadístico.

1. Planteamiento de hipótesis

Para emitir el juicio de valor se deben plantear las hipótesis, teniendo en cuenta que el enunciado del estudio pueda ser calificado de verdadero o falso. Para ello, por convención se plantean la hipótesis nula y la hipótesis alterna. Ambas, desde el punto de vista de la decisión, son excluyentes (cuando se rechaza H_0, no se rechaza H_1). En una prueba de hipótesis se compara una afirmación nula (H_0), con una alterna (H_1).

Hipótesis nula. Características

1. Se denota por H_0.
2. Es la hipótesis de trabajo (hipótesis teórica) ya que es la hipótesis que se contrasta.

3. Contiene siempre un símbolo matemático de igualdad.
4. En su formulación simbólica (Estadística) se emplea el operador igual a (=), aunque en ocasiones también se utilizan los operadores menor o igual a (≤) o mayor o igual a (≥).

Hipótesis alterna. Características

1. Se denota por H_1.
2. Es la hipótesis del investigador, debido a que es la que se necesita comprobar.
3. Contiene siempre un símbolo matemático de desigualdad.
4. Se acepta implícitamente al rechazar H_0, constituye su complemento (operadores matemáticos de desigualdad (≠, >, <).

Tipos de hipótesis

1. **Unilateral a la izquierda:** en la hipótesis alterna aparece el signo < (la región de rechazo de H_0 se encuentra a la izquierda de distribución de datos, la cola es a la izquierda). Para un α=0,05 e infinitos (∞) grados de libertad el valor critico o punto crítico es -1,65 (anexo 1) (Figura 79).

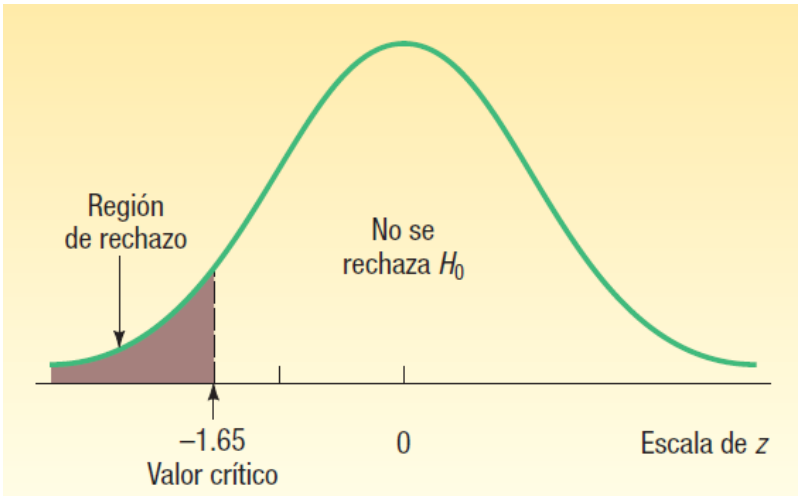


Figura 79.Distribución muestral del estadístico z; prueba de una cola a la izquierda; nivel de significancia de 0,05. Fuente: Lind et al. (2012).

2. **Unilateral a la derecha:** en la hipótesis alterna aparece el signo > (la región de rechazo de H_0 se encuentra a la derecha) Para un α=0,05 e infinitos (∞) grados de libertad el valor critico o punto crítico es 1,65 (anexo 1) (Figura 80).

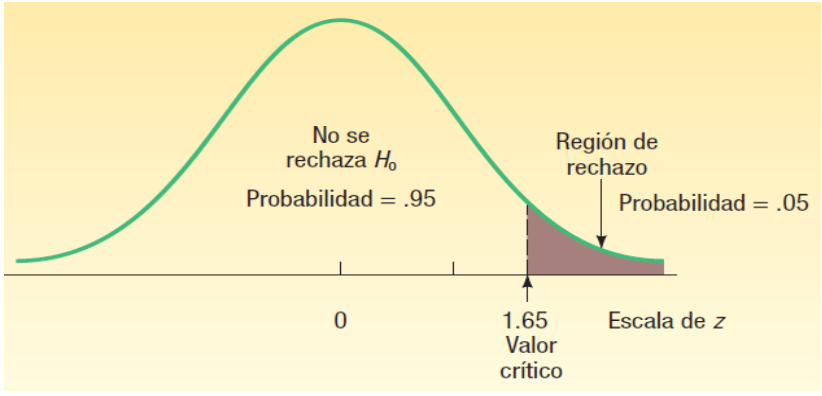


Figura 80.Distribución muestral del estadístico z; prueba de una cola a la derecha; nivel de significancia de 0,05. Fuente: Lind et al. (2012).

3. **Bilateral:** En la hipótesis alterna aparecen el signo > (la región de rechazo de H_0 se encuentra a la derecha), el signo ≠, o el signo < (la región de rechazo de H_0 se encuentra a la izquierda). Para un α=0,05 e infinitos (∞) grados de libertad el valor critico o punto crítico para ambas colas de la distribución es -1,96 y 1,96, ya que como son dos colas el α (0,05) se divide para dos (0,05/2=0,025) (anexo 1) (Figura 2.2.5).

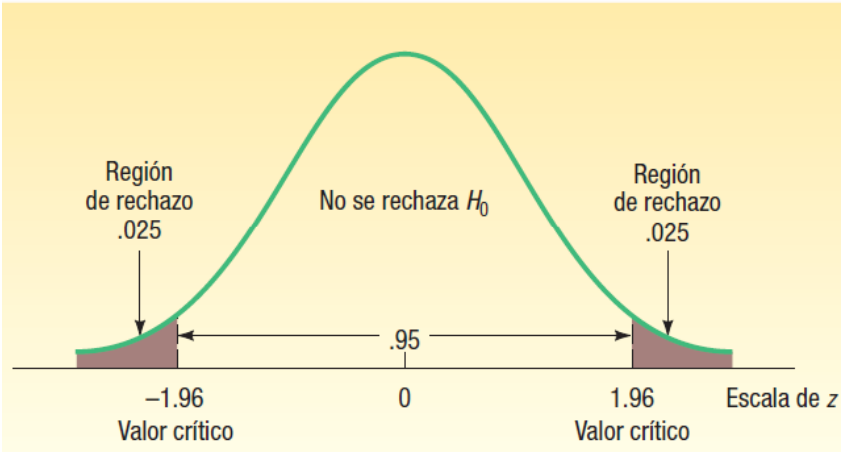


Figura 81. Regiones de aceptación y rechazo de una prueba de dos colas con un nivel de significancia de 0,05. Fuente: Lind et al. (2012).

La forma de H_1 determina el tipo de prueba de hipótesis y determina la direccionalidad de la cola, o sea, donde se ubica la región crítica o punto crítico en la curva de la distribución. Las hipótesis se plantean de forma estadística (Tabla 2.2.1) y de forma gramatical, que es la forma que se utiliza en diseño experimental.

Tabla 17. Formas de escribir las hipótesis con lenguaje estadístico cuando se dispone de dos o más tratamientos o grupos.

Tratamientos o grupos	Tipo	Hipótesis planteadas en lenguaje estadístico		
		Unilateral a la izquierda	**Unilateral a la derecha**	**Bilateral**
Dos	Igualdad	H_0: $\mu_1 \geq \mu_2$	H_0: $\mu_1 \leq \mu_2$	H_0: $\mu_1=\mu_2$
	Diferencia	H_1: $\mu_1 < \mu_2$	H_1: $\mu_1 > \mu_2$	H_1: $\mu_1 \neq \mu_2$
Tres o más	Igualdad			H_1: $\mu_1=\mu_2=\mu_3=\mu_i=\mu$
	Diferencia			H_1: $\mu_i \neq \mu$

Las hipótesis unilaterales, para el caso de disponer de tres o más tratamientos o grupos, se aplican en las comparaciones post hoc (pruebas a posteriori), cuando se quiere probar si la media de un tratamiento o grupo es mayor o menor que otro tratamiento o grupo en una dirección específica.

Hipótesis escrita en lenguaje estadístico

Para la escritura del planteamiento de hipótesis desde el punto de vista gramatical se recomienda utilizar los siguientes tips que posibilitan definir correctamente la intención del investigador (aunque existen otras formas de realizar la escritura).

Secuencia de pasos para realizar el planteamiento de hipótesis con lenguaje gramatical

1. Indicar el parámetro (la media de la POE si la variable es numérica o la proporción del parámetro si la variable es categórica).
2. Indicar cual es la variable dependiente (la numérica es medida con un instrumento mecánico que obtiene magnitudes físicas y la categoría se mide con un instrumento documental que genera dimensiones lógicas).
3. Indicar cual es el factor de estudio (engloba a todos los tratamientos o grupos).
4. Cuando se plantea H_0 siempre la oración debe indicar igualdad. Cuando se plantea H_1 siempre la oración debe indicar diferencias.

Ejemplo 1. Para el caso de la comparación de dos tratamientos o grupos.

Contexto: Un investigador tiene la hipótesis de que, en la granja Santa Inés de la UTMACH, bajo las mismas condiciones de manejo agronómico y suelo, el rendimiento agrícola (t ha^{-1}) del cultivar de soya INIAP-310 es mejor que el rendimiento agrícola del cultivar INIAP-308.

1. Planteamiento de hipótesis

- H_0: La media de rendimiento agrícola del cultivar INIAP-310 es menor o igual que la media de rendimiento agrícola del cultivar INIAP-308.
- H_1: La media de rendimiento agrícola del cultivar INIAP-310 es mayor que la media de rendimiento agrícola del cultivar INIAP-308.

Ejemplo 2. Para el caso de la comparación de más de dos tratamientos o grupos.

Contexto: Un investigador tiene la hipótesis de que, en la granja Santa Inés de la UTMACH, bajo las mismas condiciones de manejo agronómico y suelo, el rendimiento agrícola (t ha^{-1}) de tres cultivares de soya (INIAP-305, INIAP-308 e INIAP-310) es diferente.

1. Planteamiento de hipótesis

- H_0: Las medias de rendimiento agrícola de los cultivares de soya (INIAP-305, INIAP-308 e INIAP-310) son iguales.
- H_1: Al menos una media de rendimiento agrícola en los cultivares de soya (INIAP-305, INIAP-308 e INIAP-310) es diferente.

2. Establecimiento del nivel de significación de la prueba estadística

El nivel de significación estadística equivale a la máxima cantidad (magnitud) de error (Tipo I) que el investigador está dispuesto a aceptar para dar por válida la hipótesis planteada (Fisher, 1950).

Se denota por la letra del alfabeto griego α y es la probabilidad máxima que un investigador está dispuesto a aceptar para cometer un error Tipo I, es decir, rechazar incorrectamente la hipótesis nula cuando es verdadera. Alfa representa la probabilidad de cometer un error de tipo I.

El complemento del nivel de significación se denomina nivel de confiabilidad en la estimación.

Ejemplos

- Si la confiabilidad de la estimación es 93,4 %, el nivel de significación (NS=6,6 %) y α=0,066.
- Si α=0,023, el NS=2,3 % y la confiabilidad en la estimación es 97,7 %.

El nivel de significación de la prueba estadística es fijado o establecido por el investigador previamente en correspondencia con la gravedad de la consecuencia de cometer un error Tipo I.

Consideraciones a tener en cuenta para elegir a alfa

- En proyectos de investigación del área agropecuaria (α=0,05).
- Para el aseguramiento de la calidad (α=0,01).
- En encuestas en cualquier área del conocimiento (α=0,1).

Es importante advertir que los errores I y II constituyen eventos de espacios muestrales diferentes definidos por la condición de verdadera o falsa de H_0. En consecuencia, si se asignan a ellos probabilidades de ocurrencia α y β para el error tipo I y II respectivamente, será incorrecto decir que la probabilidad de error total en una prueba de hipótesis es la suma α + β.

Si se rechaza H_0, el error Tipo I se comete cuando H_0 es verdadera. Cometer un error Tipo I que consiste en rechazar la hipótesis nula siendo verdadera. Si no se rechaza H_0, siendo falsa, entonces se comete el error Tipo II. El error tipo II consiste en no rechazar la hipótesis nula siendo falsa (Tabla 18).

Tabla 18. Situaciones posibles asociadas a las distintas decisiones del investigador al desarrollar un procedimiento de prueba de hipótesis.

Decisión del investigador respecto a H_0	**Situaciones posibles**	
	H_0 falsa	**H_0 verdadera**
Rechazar H_0	Decisión correcta (probabilidad=1- β=potencia de la prueba)	Error tipo I (probabilidad=α)
No rechazar H_0	Error tipo II (probabilidad= β)	Decisión correcta (probabilidad=1-α)

El valor de α determina el punto de quiebre para tomar la decisión de rechazar o no rechazar la hipótesis nula (H_0).

3. Selección de la prueba estadística y ejecución

La elección correcta entre pruebas paramétricas y no paramétricas es crucial debido a que garantiza la validez y confiabilidad de los resultados de una investigación científica. Las pruebas paramétricas se basan en supuestos estrictos: independencia de observaciones, normalidad de los datos y homogeneidad de varianzas. Son más potentes y robustas estadísticamente cuando los supuestos se cumplen, ya que aprovechan mejor los atributos de medición de los datos.

Las pruebas paramétricas se utilizan cuando:

- Una variable es categórica y la otra es numérica que cumple los supuestos o requisitos exigidos por el modelo paramétrico.
- Las dos variables son numéricas que cumplen con los supuestos.

Las pruebas no paramétricas no requieren que los datos sigan una distribución similar a la normal. Se le denomina distribución libre. Son útiles para muestras pequeñas, datos procedentes de variables ordinales o cuando los supuestos paramétricos no se cumplen. Aunque menos potentes, ofrecen resultados más robustos en condiciones adversas y son confiables.

Las pruebas no paramétricas se utilizan cuando:

- La variable numérica no cumple, con al menos, un requisito exigido por el modelo paramétrico. Una variable es categórica y la otra es numérica, la cual no cumple algún requisito. Las dos variables son numéricas, donde al menos una no cumple con algún supuesto o requisito exigido por el modelo paramétrico.
- Las dos variables son categóricas.

Realizar una inadecuada selección genera riesgos, entre los que se encuentran:

- **Resultados inválidos:** aplicar una prueba paramétrica sin cumplir los supuestos puede llevar a conclusiones erróneas.
- **Pérdida de potencia estadística:** usar una prueba no paramétrica cuando los datos cumplen los supuestos paramétricos puede reducir la capacidad de detectar diferencias reales.
- **Interpretaciones sesgadas**: una elección incorrecta afecta la credibilidad del estudio y la repetibilidad de los hallazgos.

Para seleccionar la prueba estadística a utilizar se deben considerar varios criterios, dentro de ellos, el primero es el nivel de la investigación

científica en que se encuentre el estudio (Figura 1.2), los objetivos estadísticos, las escalas de medición de las variables y el comportamiento de los datos (Tabla 2.2.3).

Tabla 19. Principales pruebas paramétricas y no paramétricas.

Objetivo estadístico	Tipo de variable	Comportamiento de los datos	Prueba recomendada
Comparar medias entre un tratamiento o grupo y un valor de prueba, medias de dos tratamientos o grupos, o medias de un tratamiento o grupo medido antes y después.	Cuantitativa continua	Datos normales	t de Student para un tratamiento o grupos
		Datos normales, varianzas homogéneas.	t de Student para tratamientos o grupos independientes.
		Datos normales, muestras relacionadas	t de Student para medidas repetidas (pareada).
		Datos normales y varianzas no homogéneas	t de Welch.
		Datos no normales y varianzas no homogéneas	U de Mann–Whitney, Wilcoxon.
Comparar medias entre tres o más tratamientos o	Cuantitativa continua	Normalidad de datos, homogeneidad de varianzas	ANOVA de un factor intertratamientos o grupos.

grupos.			
Comparar medias con tres o más mediciones en las mismas unidades		Normalidad de datos, medidas relacionadas	ANOVA de un factor intratratamientos o medidas repetidas.
		Datos no normales y varianzas no homogéneas	H de Kruskal–Wallis (independientes), Friedman (repetidas)
Comparar proporciones entre dos grupos	Categórica dicotómica	n ≥ 30 en cada celda	Prueba z de proporciones, Ji-cuadrado 2×2.
		n pequeño (< 5)	Prueba exacta de Fisher.
Comparar proporciones entre tres o más grupos	Categórica	Frecuencias esperadas ≥ 5	Ji-cuadrado de independencia.
		Frecuencias bajas	Exacta de Fisher extendida.
Relación entre dos variables	Cuantitativas continuas	Ambas variables con datos normales	Correlación de Pearson.
		Datos no normales	Correlación de Spearman.
	Categóricas ordinales		Tau b de Kendall, Tau c de Kendall.
Regresión y predicción	Cuantitativa dependiente	Residuos normales, linealidad	Regresión lineal simple/múltiple.

		No normal, no lineal	Regresión no paramétrica (logística).
Normalidad de datos	Cuantitativa		Shapiro–Wilk, Anderson–Darling.
Varianzas homogéneas	Cuantitativa		Levene, Bartlett.
Independencia de observaciones	Categórica		Ji-cuadrado de independencia.
Bondad de ajuste	Categórica		Ji-cuadrado de bondad de ajuste.

4. Lectura del p-valor

En la actualidad con el uso del software estadístico se utiliza el p-valor para tomar la decisión de rechazar o no rechazar H_0. El p-valor es la cuantificación (cantidad) del error que se comete realmente al realizar la prueba estadística (al ser un valor de probabilidad varía entre 0 y 1 (Figura 82). Según Triola (2018) el p-valor se describe como la probabilidad de cometer un error tipo I si se rechaza la hipótesis nula.

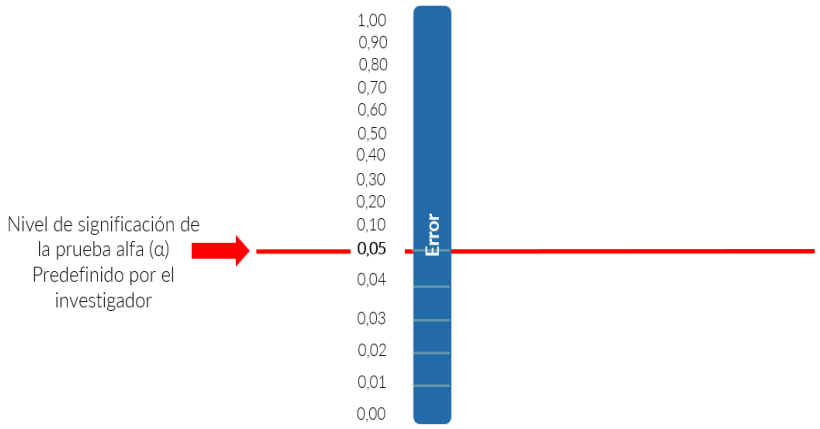


Figura 82. Representación gráfica del nivel de significación (α) en pruebas de hipótesis, indicando el umbral de error tipo I definido previamente por el investigador (α = 0,05).

En caso de presentar una prueba de hipótesis de una sola cola (unilateral) el p-valor obtenido en la prueba estadística realizada se debe dividir entre dos; y es el valor que se toma en cuenta para realizar la lectura del p-valor y tomar la decisión.

Cuando se realiza el contraste de hipótesis de forma manual se obtienen los estadísticos calculados mediante algoritmos y después con el uso de las tablas se obtienen los puntos críticos para efectuar la comparación.

Calcular el estadístico t

Para desarrollar el procedimiento de calcular el estadístico t se debe utilizar la fórmula establecida para la comparación de tratamientos o grupos independientes.

$$t_{cal} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S^2{}_1}{n_1} + \frac{S^2{}_2}{n_2}}}$$

Dónde:

$\bar{X}_1$ y $\bar{X}_2$ son las medias muestrales de los dos grupos (fincas bananeras).

S_1^2 y S_2^2 son las varianzas muestrales de cada finca.

n_1 y n_2 son los tamaños de las muestras de cada finca.

En la prueba t de Student para tratamientos o grupos independientes los grados de libertad (gl) se calculan de la siguiente manera: gl=n_1+ n_2 – 2.

Calcular el estadístico F

Para obtener el estadístico F se deben calcular todos los componentes de la tabla ANOVA de un factor intertratamientos o grupos (Tabla 2.2.4).

Fuente de variación	Suma de cuadrados	gl	Cuadrados medios	Estadístico $F_{calculado}$	$F_{tabulado}$
Tratamientos o grupos (varianza del modelo)	$SCTr = n_i \sum_{i=1}^{k} (\bar{X}_{i.} - \bar{X}_{i..})^2$	k-1	$CMTr = \frac{SCTr}{k-1}$	$F_{cal} = \frac{SCTr}{SCE}$	α seleccionado gl numerador gl denominador
Errores (varianza residual)	$SCE = \sum_{i=1}^{k} \sum_{j=1}^{n} (Y_{ij} - \bar{X}_{i.})^2$	n-k	$CME = \frac{SCE}{n-k}$		
Total	$SCT = \sum_{i=1}^{k} \sum_{j=1}^{n} (Y_{ij} - \bar{X}_{..})^2$	**n-1**			

Figura 83. Esquema de la tabla del análisis de varianza de un factor intertratamientos o grupos.

Nota: k=cantidad de tratamientos o grupos. n=tamaño de la muestra (número todas las observaciones realizadas).

Para calcular los grados de libertad en un ANOVA de un factor intertratamientos o grupos se debe considerar lo siguiente:

- gl intertratamientos o grupos (k-1). gl del numerador.
- gl intratratamientos o grupos (n-k). gl del denominador.
- gl total (n-1).

5. Tomar la decisión

Cuando el procedimiento se desarrolla con el uso de un software estadístico la decisión se encuentra relacionada con el p-valor calculado en la prueba estadística (error cometido realmente) y alfa, o sea, el nivel de significación fijado por el investigador previamente (error planificado).

- Si el p-valor obtenido en la prueba estadística es ≤ α planificado se rechaza H_0.
- Si el p-valor obtenido en la prueba estadística es > α planificado no se rechaza H_0.

Regla de decisión en la prueba t de Student para tratamientos o grupos independientes (sin el uso de software estadístico).

Para una prueba bilateral (dos colas):

- Si $t_{cal} \geq t_{\alpha/2}$ (punto crítico) se rechaza H_0.
- Si $t_{cal} < t_{\alpha/2}$ (punto crítico) no se rechaza H_0.

Para una prueba unilateral (cola a la derecha):

- Si $t_{cal} \geq t_{\alpha}$ (punto crítico) se rechaza H_0.
- Si $t_{cal} < t_{\alpha}$ (punto crítico) no se rechaza H_0.

Para una prueba unilateral (cola a la izquierda):

- Si $t_{cal} \leq -t_{\alpha}$ (punto crítico) se rechaza H_0.
- Si $t_{cal} > -t_{\alpha}$ (punto crítico) no se rechaza H_0.

Regla de decisión en la prueba ANOVA de un factor intertratamientos o grupos (sin el uso de software estadístico).

Para tomar la decisión cuando el procedimiento se realiza de forma manual se debe comparar el valor del $F_{calculado}$ con el valor $F_{tabulado}$ o $F_{crítico}$.

- Si $F_{calculado} \geq F_{crítico}$: se rechaza la hipótesis nula (se presentan diferencias estadísticas significativas entre los tratamientos o grupos).
- Si $F_{calculado} < F_{crítico}$: no se rechaza la hipótesis nula (no se presenta evidencia suficiente de diferencias entre tratamientos o grupos.

En la práctica el caso de igualdad de F_{cal} y F_{tab} es muy poco probable que ocurra debido a que los valores críticos son números con varios decimales y los estadísticos (t o F) calculados, raramente coinciden exactamente. Sin embargo, teóricamente se interpreta como rechazo de la hipótesis nula.

A continuación, se detallan los principales elementos teóricos y prácticos (procedimiento desarrollado de forma manual sin la utilización del software estadístico) relacionados con las pruebas paramétricas utilizadas con mayor frecuencia en la experimentación agropecuaria.

Test t (t de Student)

Procedimiento estadístico utilizado, principalmente, para realizar todas las comparaciones posibles, entre dos tratamientos o grupos independientes o relacionados en función de una variable medida (variable dependiente numérica); por ello, para su desarrollo de debe desarrollar el procedimiento de prueba de hipótesis.

Distribución t de Student. Surgimiento

William Sealy Gosset (1876-1937) fue un estadístico, nacido en Canterbury, Inglaterra. Tras graduarse en 1899, se incorporó a las destilerías Guinness en Dublín donde aplicó sus conocimientos

estadísticos (desarrolló estudios de prueba y error) y trabajó con Karl Pearson. Otro investigador de Guinness había publicado anteriormente un artículo que contenía secretos industriales de la destilería, por lo que, se prohibió la publicación de artículos. Por esta situación Gosset no podía publicar sus trabajos usando su propio nombre.

Gosset en 1908 (32 años), publicó el artículo "El error probable de una media" en la revista Biometrika (creada por Karl Pearson), no con su nombre, sino que publicó todos sus artículos usando el pseudónimo Student, aunque fue Ronald Aylmer Fisher (1890-1962) quien apreció la importancia de los trabajos de Gosset sobre muestras pequeñas.

Gosset fue amigo tanto de Pearson como de Fisher, lo que constituye un logro, ya que ambos nunca lo hicieron entre ellos. Sin embargo, su logro más famoso se conoce ahora como la distribución t de Student, que de otra manera hubiera sido la distribución t de Gosset.

El test t puede utilizarse para:

- Determinar si la media de un único tratamiento o grupo difiere de un valor de prueba conocido (prueba t de Student de un tratamiento o grupo). Se utiliza cuando se desarrolla un preexperimento con un tratamiento o se estudia un solo grupo y se efectúan las mediciones de la variable numérica; y se compara con un valor de prueba teórico o establecido previamente a través de la investigación documental. Los datos deben cumplir con los requisitos del modelo paramétrico.
- Determinar si dos tratamientos o grupos difieren entre sí en función de una variable numérica que cumple los supuestos

(prueba t de Student para tratamientos o grupos independientes). Se utiliza cuando se desarrolla un experimento con dos tratamientos o se estudian dos grupos independientes y la variable medida es numérica y cumple con los requisitos del modelo paramétrico. Es la prueba estadística más utilizada en investigación científica cuando se tienen dos tratamientos o grupos.

- Determinar si se presenta una diferencia significativa en medidas pareadas de una variable medida antes y otra después (prueba t de Student para tratamientos o grupos pareados, también se le conoce como t de Student para medidas repetidas). Se utiliza cuando se desarrolla un preexperimento con un tratamiento o se estudia un solo grupo y se efectúan las mediciones de la variable numérica antes y después de ocurrida una intervención o un acontecimiento. La diferencia debe cumplir con los requisitos del modelo paramétrico.

t de Student para tratamientos o grupos independientes

Técnicamente se puede describir a la prueba t de Student para tratamientos o grupos independientes como aquella prueba estadística que se utiliza en un modelo en el que una variable explicativa (variable independiente categórica dicotómica), intenta explicar una variable respuesta (variable dependiente aleatoria) numérica que cumple con los requisitos exigidos por el modelo paramétrico). Esta situación se plantea cuando se comparan dos tratamientos o grupos en cuanto a una variable de estudio numérica. Es una prueba de contraste de hipótesis sobre los

valores de dos medias obtenidas en dos tratamientos o grupos o independientes.

Propiedades

1. Es una distribución de probabilidad que se utiliza cuando se quiere estimar la media de una población distribuida según una normal, cuando el tamaño de la muestra utilizada para la estimación es pequeño (30 o menos observaciones) y la varianza de la POE es desconocida.

2. Tiene propiedades parecidas a la distribución normal: media cero y simétricas con respecto a la media. Se diferencia en cuando a los grados de libertad; la distribución normal se utiliza para muestras grandes (más de 30 observaciones o casos).

3. La gráfica de la prueba t varía hasta 30, sin embargo, cuando pasa de 30 se asemeja a la distribución Z, por eso, cuando los datos pasan de 30 el valor crítico de t y z son iguales, se hace igual a la N (0, 1) (la varianza decrece hasta 1 cuando los gl pasan de 30). Elemento clave en experimentación que sustenta el criterio de recolectar más de 30 datos por cada tratamiento en un experimento.

4. Para su utilización no es obligatorio que los tamaños de los grupos sean iguales, ni tampoco es necesario conocer la dispersión de los dos grupos.

5. Es un procedimiento paramétrico que requiere el cumplimiento de requisitos o supuestos para utilizarse.

Cuando se aplica la prueba t de Student para grupos independientes

- Cuando se tienen dos tratamientos o grupos independientes, con igual o diferente cantidad de observaciones y las varianzas son desconocidas.
- Cuando la variable numérica en cada tratamiento o grupo cumple los supuestos o requisitos de independencia de observaciones, normalidad de datos y homogeneidad de varianzas (asunción estadística).

Asunción estadística

Condición teórica que debe cumplirse para validar un procedimiento estadístico paramétrico (independencia de observaciones, normalidad de datos y homogeneidad de varianzas), sea el test t o el test F. En ciencias agropecuarias es frecuente verificar estas condiciones mediante pruebas estadísticas y gráficos.

1. Independencia de observaciones

El supuesto de independencia de observaciones, es una condición fundamental en varios métodos estadísticos y modelos de inferencia. Significa que cada dato u observación del conjunto no debe estar influenciado por otra; o sea, el valor o dato obtenido en una observación no depende del valor de la otra observación (cada dato debe ser recogido sin que dependa del otro). Se deben obtener puntuaciones diferentes debido a que provienen de objetos o sujetos ubicados en grupos o unidades experimentales diferentes, no relacionados o influenciados

unos con otros. Cada observación aporta información única y no está relacionada con las demás.

Ejemplo: si lanzas una moneda varias veces, cada lanzamiento es independiente del anterior. El resultado de un lanzamiento no afecta al siguiente.

Elementos importantes a considerar

1. Las puntuaciones o valores de una muestra no pueden ser similares a la otra muestra a priori, debido a que las unidades de estudio o de muestreo son distintas en cada tratamiento o grupo.
2. Las unidades de estudio o de muestreo de un tratamiento o un grupo no pueden influir en los datos obtenidos en el otro tratamiento o grupo, ya que sus características son distintas.

Cómo se garantiza la independencia de observaciones

1. En la aleatorización de los tratamientos en un experimento o selección aleatoria de los grupos en un estudio observacional.
2. En la recolección de datos cada unidad de estudio o de muestreo se selecciona de manera independiente (se aplica el muestreo aleatorio simple, donde cada una tenga la misma probabilidad de formar parte del estudio (diseño de muestreo y recolección de datos). Se debe considerar eliminar el efecto de borde en parcelas experimentales, lo cual consiste en no realizar recolección de datos en las unidades de estudio que puedan estar influenciadas por factores externos.

La independencia de observaciones es un requisito importante debido a que varios métodos estadísticos (t de Student, análisis de varianza, regresión lineal) asumen independencia de observaciones o residuos, para que las fórmulas de la varianza, error típico o estándar de la media y significancia sean válidas.

Cuando se presentan observaciones relacionadas o dependientes (por ejemplo, medidas repetidas de la misma persona, animal o planta, datos de series temporales, o encuestas en grupos homogéneos), los resultados pueden ser engañosos: se subestima la variabilidad y se sobreestima la significancia estadística.

2. Normalidad de datos

La distribución normal de los datos es un modelo probabilístico continuo que describe muchas variables biológicas. En análisis experimental se asume que los residuos siguen esta distribución para validar la prueba paramétrica. Plantea que los valores de la variable de estudio dentro del tratamiento o grupo o en cada grupo (dos o más) o la diferencia entre las medidas antes y después de efectuar las mediciones se debe presentar una distribución similar a la distribución normal (distribución teórica de referencia), denominada como distribución de Gauss o distribución gaussiana.

Propiedades de la Distribución de Gauss o Distribución gaussiana

- **La media aritmética (promedio) es central;** o sea, se ubica en el centro de la distribución de datos; lo que significa que la mayor

parte de los datos se concentran alrededor de ese valor. La media coincide con la mediana y la moda.

- **Es unimodal.** La distribución tiene un solo "pico" o máximo, que corresponde al valor más frecuente (la moda), que coincide con la media. En distribuciones multimodales se presentan múltiples picos.

- **Es simétrica respecto a la media aritmética.** Los valores a la izquierda y a la derecha de la media se distribuyen de manera idéntica, lo que implica que la probabilidad de observar un valor por debajo o por encima de la media es la misma, si se mide a igual distancia.

- **Es asintótica.** Los extremos de la curva se acercan al eje horizontal; sin embargo, nunca lo tocan. Aunque la probabilidad de valores muy alejados de la media es extremadamente baja, nunca es exactamente cero. En otras palabras, siempre existe una posibilidad (aunque mínima) de observar valores extremos (Figura 84) (Peter, 2026).

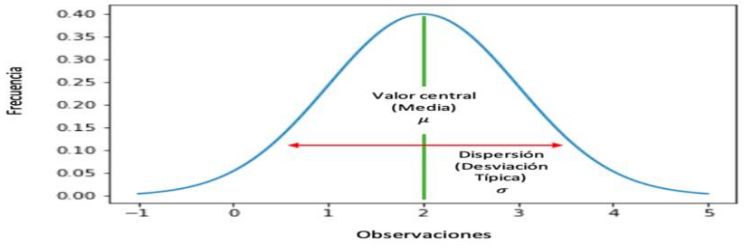


Figura 84.Representación gráfica de las propiedades de la distribución normal.

Fuente: LibreTexts (2026).

- **El 95,44 % de los datos se encuentran a dos desviaciones estándar de la media aritmética**. El 68,26 % de los datos se encuentran a una desviación estándar de la media aritmética y según la regla empírica el 99,74 % de los datos se encuentran a tres desviaciones estándar de la media (Figura 2.2.8) (Peter, 2026).

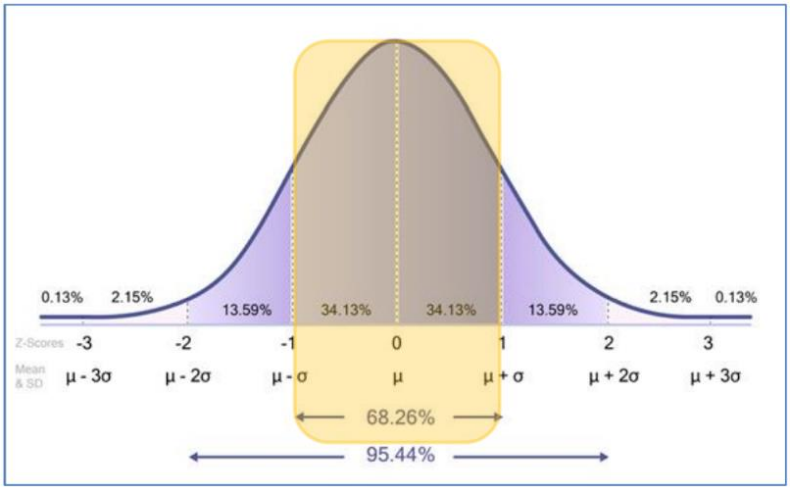


Figura 85. Comportamiento de la variabilidad de los valores en una distribución normal.

Fuente: LibreTexts (2026).

Elementos importantes a considerar

1. Para verificar el requisito de normalidad de datos se debe emitir un juicio de valor, por ello, para su comprobación se debe desarrollar un procedimiento de prueba de hipótesis.
2. La hipótesis teórica parte del planteamiento de hipótesis (lenguaje gramatical) de que la distribución de datos de la variable aleatoria es similar a la distribución normal.

3. La verificación estadística se realiza con los test de Shapiro–Wilk y Anderson–Darling.
4. La violación del supuesto de normalidad de datos no afecta de forma importante siempre que los tamaños muestrales sean grandes (mayores a 30 observaciones). En este caso, no se considera obligatorio el cumplimiento del requisito.

3. Homogeneidad de varianzas (Homocedasticidad)

La homogeneidad de varianzas es un supuesto o requisito fundamental, que exige que la variabilidad de los datos de un tratamiento o grupo debe ser similar a la variabilidad del otro tratamiento o grupo. Se evalúa mediante las pruebas de Levene, Bartlett o inspección gráfica.

Por otro lado, la variabilidad entre tratamientos o grupos se expresa en términos de varianza, utilizando el diagrama de caja y sesgos calculado a partir de la media aritmética. Cuando la proyección de la línea media del primer tratamiento o grupo no coincide con la caja del segundo tratamiento o grupo, se considera que la variabilidad es diferente (Figura 2.2.9A). Si la proyección de la línea media del primer tratamiento o grupo se encuentra dentro de la caja del segundo tratamiento o grupo, la variabilidad se interpreta como igual (Figura 86).

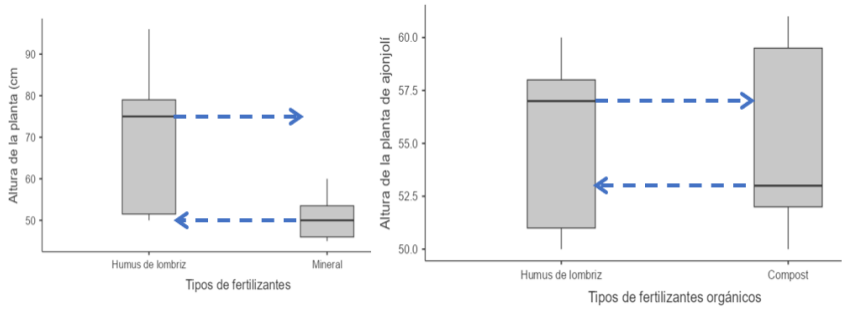


Figura 86.Verificación gráfica de la homogeneidad de varianzas entre dos tratamientos o grupos independientes. A. Varianzas diferentes. B. Varianzas iguales.

Elementos importantes a considerar

- Se asume que las varianzas de los dos tratamientos o grupos son homogéneas.
- Se evalúa la igualdad de las varianzas de una variable calculada para dos tratamientos o grupos.
- La verificación estadística se realiza con el test de Levene (Levene, 1960).
- En el caso de que el contraste de homogeneidad de varianzas resulte significativo (p-valor ≤ alfa), se procede a utilizar el Test de Welch, esto es, prueba t para dos tratamientos o grupos independientes y se supone varianzas desiguales (Gastwirth et al., 2009).
- Si no se verifica el cumplimiento de este supuesto los resultados de la prueba t de Student no tienen ninguna validez y se puede llegar a plantear conclusiones erróneas.

Alternativas a utilizar

- Transformación de los datos (raíz cuadrada, elevar al cuadrado).
- Utilizar la prueba U de Mann-Whitney (alternativa no paramétrica). La recomendada en investigación experimental.

Ejemplo de prueba t de Student para tratamientos o grupos independientes (sin el uso de Jamovi)

Contexto

En una investigación desarrollada en la UTMACH se dispone de dos grupos (Fincas bananeras del cantón Cañaquemada) y se recolectaron datos de la variable peso de racimos de banano (kg) en cinco puntos de muestreo.

Datos recolectados

Finca 1 (n_1= 10): 20,4; 22,3; 23,8; 25,6; 28,9; 27,3; 26,7; 25,9; 24,6; 28,2.

Finca 2 (n_2= 10): 24,3; 26,2; 27,7; 29,1; 32,3; 31,5; 28,6; 27,6; 26,9; 33,1.

Requerimiento del ejercicio

Verifique si se presentan o no diferencias estadísticas significativas (se debe desarrollar el procedimiento de forma manual) entre las fincas bananeras en función del peso de racimos de banano. Utilice un 95 % de confiabilidad en la estimación.

Procedimiento

1. Planteamiento de hipótesis

H_0: La media de peso de racimos de banano en la finca 1 es igual a la media de peso de racimos de banano en la finca 2.

H_1: La media de peso de racimos de banano en la finca 1 es diferente a la media de peso de racimos de banano en la finca 2.

2. Establecimiento del nivel de significación de la prueba

α=0,05. NS=5,0 %.

3. Calcular las medias muestrales y las varianzas de cada finca

Tabla 20.Calcular las medias muestrales y las varianzas de cada finca.

Fincas	Peso de racimos (kg)	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
Finca 1	20,4	-4,97	24,70
Finca 1	22,3	-3,07	9,42
Finca 1	23,8	-1,57	2,46
Finca 1	25,6	0,23	0,05
Finca 1	28,9	3,53	12,46
Finca 1	27,3	1,93	3,72
Finca 1	26,7	1,33	1,77
Finca 1	25,9	0,53	0,28
Finca 1	24,6	-0,77	0,59
Finca 1	28,2	2,83	8,01
Σ	**253,7**	**Σ**	**63,48**
$\bar{X}$	**25,37**	S^2	**7,05**
Finca 2	24,3	-4,43	19,62
Finca 2	26,2	-2,53	6,40

Finca 2	27,7	-1,03	1,06
Finca 2	29,1	0,37	0,14
Finca 2	32,3	3,57	12,74
Finca 2	31,5	2,77	7,67
Finca 2	28,6	-0,13	0,02
Finca 2	27,6	-1,13	1,28
Finca 2	26,9	-1,83	3,35
Finca 2	33,1	4,37	19,10
Σ	**287,3**	**Σ**	**71,38**
$\bar{X}$	**28,73**	S^2	**7,93**

4. Calcular el estadístico t

$$t_{cal} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S^2{}_1}{n_1} + \frac{S^2{}_2}{n_2}}}$$

$t_{cal} = \frac{25,37-28,73}{\sqrt{\frac{7,05}{10}+\frac{7,93}{10}}}$ $t_{cal} = \frac{-3,36}{\sqrt{0,71+0,79}}$ $t_{cal} = \frac{-3,36}{\sqrt{1,498}}$ $t_{cal} = \frac{-3,36}{1,22} =$ $t_{cal} = -2,75$

5. Calcular los grados de libertad (gl)

Los gl para la prueba t de Student para tratamientos o grupos independientes se calculan de la siguiente manera:

gl=n_1+n_2-2 gl=10+10-2 gl=20-2 gl=18

6. Obtener el t tabulado

Para obtener el t_{tab} o punto crítico se debe utilizar la tabla de la distribución t de Student (anexo 1), el cual depende del nivel de

significación establecido previamente por el investigador (α=0,05). Debido a que es una prueba bilateral (lo define el planteamiento de H_1) se debe dividir el alfa (0,05/2=0,025) y calcular los grados de libertad (10+10-2=18) y realizar la interpolación en la tabla para buscar el estadístico t_{tab}.

$$t_{tab}=2,10$$

7. Regla de decisión

Para este ejercicio la regla de decisión se establece de la siguiente forma:

Como el t_{cal} (-2,75) ≤ t_{tab} (-2,10) se rechaza H_0; ya que el t_{cal} cae en la región de rechazo de H_0.

8. Interpretación de los resultados

El peso de racimos de banano en la finca 2 (28,73 kg); es mayor y diferente estadísticamente al peso de racimos de banano obtenidos en la finca 1 (25,37 kg), lo que supone que en la Finca 2 se realiza un mejor manejo agronómico del cultivo o la condición edafoclimática de es más apropiada.

Limitaciones del Test t

1. Es tedioso realizar todas las posibles combinaciones (si se dispone de tres tratamientos (A, B y C) se deben realiza tres pruebas de hipótesis; o sea, A-B, A-C y B-C).

2. Es menos estable debido a que utiliza parte de la evidencia (debido a que considera dos tratamientos, siempre los restantes no se utilizan).
3. Cuando se realizan varias pruebas t de Student se acumula el error de Tipo I (si se dispone de tres grupos se cometen tres errores).

Las limitaciones que presenta el test t se solucionan mediante la utilización del test F de Snedecor, o análisis de varianza (ANOVA), que es una extensión de la prueba t de Student para tratamientos o grupos independientes; que tiene la ventaja de realizar comparaciones en estudios experimentales u observacionales con más de dos tratamientos o grupos independientes en función de la media de la variable dependiente numérica (debe cumplir todos los requisitos). sin que se incremente el error de Tipo I.

El test F analiza al conjunto de tratamientos o grupos; y parte del planteamiento de tener k tratamientos o grupos (más de dos) independientes, procedentes de poblaciones normales, con varianzas iguales (requisitos del modelo paramétrico).

¿Cómo surge el análisis de varianza?

Las técnicas iniciales de análisis de varianza fueron desarrolladas por el estadístico y genetista Ronald A. Fisher, en los años 1920 y debido a que se basa en la distribución F de Fisher como parte del contraste de hipótesis, a menudo se le llama “análisis de varianza de Fisher”

El test F incluye varios tipos de pruebas estadísticas paramétricas, entre las que se encuentran:

1. Análisis de varianza de un factor intertratamientos o grupos.
2. Análisis de varianza de un factor intratratamientos o grupos.
3. El análisis de varianza factorial intertratamientos o grupos.

ANOVA de un factor intertratamientos o grupos

En un experimento unifactorial cuando el investigador manipula un factor de estudio con más de dos tratamientos (se denominan grupos si fuese un estudio no experimental) la comparación que se realiza entre ellos, para conocer si se presentan o no diferencias estadísticamente significativas en función de la variable respuesta (numérica y debe cumplir con los requisitos del modelo paramétrico), se realiza con la prueba de hipótesis (procedimiento) de análisis de varianza de un factor intertratamientos o grupos (test F).

Definición

El ANOVA de un factor intertratamientos es una prueba estadística paramétrica que contrasta más de dos tratamientos o grupos (factor de estudio), a través de la comparación de la variabilidad inter (entre) tratamientos o grupos (la cual es debida al azar y al efecto de los tratamientos, si es que existe) y la variabilidad intra (dentro) tratamientos o grupos (la cual es debida al azar), obteniéndose el estadístico F). Permite descomponer la variación total de los datos en componentes atribuibles a los tratamientos o grupos y al error

experimental, con el fin de evaluar si se presentan o no diferencias estadísticas significativas entre las medias de los tratamientos o grupos.

Objetivo

Contrastar la hipótesis de que los efectos de los tratamientos o grupos del factor de estudio (tres o más) son nulos o iguales (H_0), versus que al menos uno no lo es (H_1).

Propósito

1. Se utiliza para conocer si se presentan o no diferencias estadísticamente significativas entre los tratamientos (grupos) objeto de estudio (más de dos) en función de la variable de estudio.

2. Es una prueba Ad hoc ya que no indica dónde se encuentran las similitudes o las diferencias; lo cual, se establece mediante pruebas *Post hoc* (Pruebas de rangos y comparaciones múltiples (Tukey, Duncan), o sea, se debe realizar una prueba a posteriori.

Cuando se utiliza

1. Cuando se dispone de un factor de estudio con más de dos tratamientos o grupos establecidos de forma independiente y se contrastan para conocer si sus medias son iguales o diferentes estadísticamente.

2. Cuando la variable de estudio es numérica y cumple con todos los supuestos para realizar pruebas paramétricas

Requisitos que se deben cumplir para utilizar el modelo ANOVA paramétrico

- Variable numérica (por defecto).
- Independencia de observaciones o de errores (se garantiza con la aleatorización de los tratamientos). No se verifica estadísticamente.
- Normalidad de datos (se verifica con el test estadístico de Shapiro-Wilk).
- Homogeneidad de varianzas (se verifica con el test estadístico de Levene). La variabilidad de las observaciones debe ser homogénea dentro de cada grupo.

Precisiones

- No es necesario que el tamaño de la muestra de los tratamientos o grupos sea el mismo.
- Cuanto mayor es el tamaño de los tratamientos o grupos (mayores de 30 observaciones es menos importante garantizar el cumplimiento del supuesto de normalidad de datos)

Ejemplo de ANOVA de un factor intertratamientos o grupos (sin el uso de Jamovi)

Contexto

Un profesor universitario que imparte la catedra de Estadística quiere probar si el método de enseñanza aprendizaje influye en las calificaciones de los estudiantes. Para ello, selecciona tres métodos de enseñanza-aprendizaje diferentes (A, B, C) y selecciona aleatoriamente

15 estudiantes, los cuales se ubican en tres grupos de cinco estudiantes y se les enseña un método diferente.

Datos recolectados

Resultados de las puntuaciones obtenidas por cada estudiante en el examen realizado (10 puntos).

Tabla 21.Resultados de las puntuaciones.

No.	Método de EA	Puntos obtenidos	No.	Método de EA	Puntos obtenidos	No.	Método de EA	Puntos obtenidos
1	A	3	6	B	5	11	C	7
2	A	2	7	B	2	12	C	4
3	A	1	8	B	4	13	C	5
4	A	1	9	B	2	14	C	3
5	A	4	10	B	3	15	C	6

Requerimiento del ejercicio

Verifique estadísticamente, mediante el desarrollo del procedimiento manual, si se presenta o no diferencias estadísticas entre los métodos de enseñanza aprendizaje en función de las puntuaciones obtenidas por los estudiantes en el examen de Estadística. Utilice un 95 % de confiabilidad en la estimación.

Procedimiento estadístico

1. Planteamiento de hipótesis

H_0: Las medias de las puntuaciones obtenidas por los estudiantes de Estadística con la utilización de los métodos de enseñanza aprendizaje son iguales.

H_1: Al menos una media de las puntuaciones obtenidas por los estudiantes de Estadística con la utilización de los métodos de enseñanza aprendizaje es diferente.

2. Establecimiento del nivel de significación de la prueba

α=0,05. NS=5,0 %.

3. Construcción del gráfico cartesiano

El procedimiento para el desarrollo del ejercicio sin el uso del software estadístico Jamovi continua con la construcción del gráfico cartesiano que representa las fuentes de variabilidad en un ANOVA de un factor intertratamientos o grupos considerando los siguientes pasos:

- Representar las puntuaciones mediante la interpolación en el espacio bidimensional de las calificaciones obtenidas en el examen por cada uno de los participantes.
- Calcular la media general o gran media ($\overline{X}_{..}$=3,47 puntos) y ubicar una línea horizontal discontinua de color rojo que la represente en el gráfico.

- Calcular la media de cada tratamiento o grupo ($\overline{X}_{.1} = 2{,}20\ puntos,\ \overline{X}_{.2} = 3{,}20\ puntos, \overline{X}_{.3} = 5{,}50\ puntos,$); y ubicar una línea continua horizontal del color que representa a dicho tratamiento o grupo.
- Representar los residuos intertratamientos o grupos (entre) con una línea continua vertical de color celeste.
- Representar los residuos intratratamientos o grupos (dentro) con una línea continua vertical del color que identifica a los tratamientos o grupos (Figura 87).

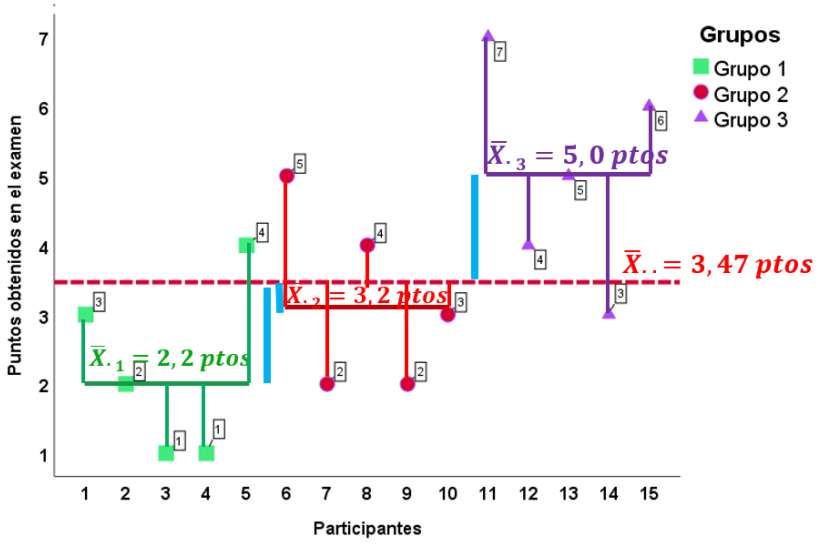


Figura 87.Gráfico cartesiano que representa las fuentes de variabilidad en un ANOVA de un factor intertratamientos o grupos.

4. Calcular las sumas de cuadrados

El análisis de varianza se basa en la partición o descomposición de la variabilidad total asociada al experimento (suma de cuadrados totales) en dos estimaciones independientes de varianza: La varianza entre los

tratamientos experimentales o grupos y la varianza dentro de los tratamientos experimentales o grupos.

Suma de cuadrados intertratamientos o grupos (SCTr)

Se refiere a la variabilidad explicada por las diferencias entre las medias de los tratamientos o grupos que forman el factor de estudio. Esta estimación mide cuan variable es la media de cada tratamiento con respecto a la media general.

$$SCTr = n_i \sum_{i=1}^{k} (\bar{X}_{i.} - \bar{X}_{i..})^2$$

$$SCTr = 5\ (2{,}20 - 3{,}47)^2 + (3{,}20 - 3{,}47)^2 + (5{,}00 - 3{,}47)^2$$

$$SCTr = 5\ (-1{,}27)^2 + (-0{,}27)^2 + (1{,}53)^2$$

$$SCTr = 5\ (1{,}61 + 0{,}07 + 2{,}35)$$

$$SCTr = 5\ (4{,}03) = 20{,}15$$

Sumas de cuadrados de los errores (SCE)

Se refiere a la variabilidad existente dentro de cada tratamiento (residuos). Es la variabilidad no explicada por las diferencias entre las medias de los tratamientos o grupos.

$$SCE = \sum_{i=1}^{k} \sum_{j=1}^{n} (Y_{ij} - \bar{X}_{i.})^2$$

$$SCE = (3 - 2{,}2)^2 + (2 - 2{,}2)^2 + (1 - 2{,}2)^2 + (1 - 2{,}2)^2 + (4 - 2{,}2)^2 + (5 - 3{,}2)^2 + (2 - 3{,}2)^2 + (4 - 3{,}2)^2 + (2 - 3{,}2)^2 + (3 - 3{,}2)^2 + (7 - 5)^2 + (4 - 5)^2 + (5 - 5)^2 + (3 - 5)^2 + (6 - 5)^2$$

$$SCE = (0{,}8)^2 + (-0{,}2)^2 + (-1{,}2)^2 + (-1{,}2)^2 + (1{,}8)^2 + (1{,}8)^2 + (-1{,}2)^2 + (0{,}8)^2 + (-1{,}2)^2 + (-0{,}2)^2 + (2)^2 + (-1)^2 + (0)^2 + (-2)^2 + (1)^2$$

$$SCE = 0{,}64 + 0{,}04 + 1{,}44 + 1{,}44 + 3{,}24 + 3{,}24 + 1{,}44 + 0{,}64 + 1{,}44 + 0{,}04 + 4 + 1 + 0 + 4 + 1$$

$$SCE = 23{,}60$$

Suma de cuadrados totales (SCT)

Se refiere a la variabilidad total en los datos; o sea, a la variación de cada observación respecto a la media general. La variación total (SCT) es igual a la suma de la variación entre tratamientos o grupos (SCTr) y la variación intratratamientos o grupos (SCE).

$$SCT = \sum_{i=1}^{k} \sum_{j=1}^{n} \left(Y_{ij} - \bar{X}_{..}\right)^2$$

$$SCE = (3 - 3{,}47)^2 + (2 - 3{,}47)^2 + (1 - 3{,}47)^2 + (1 - 3{,}47)^2 + (4 - 3{,}47)^2 + (5 - 3{,}47)^2 + (2 - 3{,}47)^2 + (4 - 3{,}47)^2 + (2 - 3{,}47)^2 + (3 - 3{,}47)^2 + (7 - 3{,}47)^2 + (4 - 3{,}47)^2 + (5 - 3{,}47)^2 + (3 - 3{,}47)^2 + (6 - 3{,}47)^2$$

$$SCE = (-0{,}47)^2 + (-1{,}47)^2 + (-2{,}47)^2 + (-2{,}47)^2 + (0{,}53)^2 + (1{,}53)^2 + (-1{,}47)^2 + (0{,}53)^2 + (-1{,}47)^2 + (-0{,}47)^2 + (3{,}53)^2 + (0{,}53)^2 + (1{,}53)^2 + (-0{,}47)^2 + (2{,}53)^2$$

$$SCE = 0{,}22 + 2{,}16 + 6{,}10 + 6{,}10 + 0{,}28 + 2{,}34 + 2{,}16 + 0{,}28 + 2{,}16 + 0{,}22 + 12{,}46 + 0{,}28 + 2{,}34 + 0{,}22 + 6{,}40$$

$$SCE = 43{,}75$$

5. Calcular los grados de libertad (gl)

gl intertratamientos o grupos (k-1): 3-1=2

gl intratratamientos o grupos (n-k): 15-3=12

gl total (n-1): 15-1=14

6. Calcular los cuadrados medios (CM)

Cuadrados medios de los tratamientos o grupos (CMTr)

$$CMTr = \frac{SCTr}{k-1} \qquad CMTr = \frac{20{,}15}{2} = 10{,}07$$

Cuadrados medios de los errores (CME)

$$CME = \frac{SCE}{n-k} \qquad CME = \frac{23{,}60}{12} = 1{,}97$$

7. Calcular el estadístico F.

$$F_{cal} = \frac{SCTr}{SCE} \qquad F_{cal} = \frac{10{,}07}{1{,}97} = 5{,}12$$

8. Regla de decisión

Para encontrar el valor del $F_{tabulado}$ o $F_{crítico}$ se utiliza la tabla de la distribución F de Snedecor (anexo 2), considerando el nivel de significancia establecido previamente (α=0,05) y los grados de libertad del numerador (2) y del denominador (14). En este caso en la tabla por aproximación se escoge el 15 para el denominador y se interpola obteniendo un $F_{tabulado}$=3,6823, el cual se utiliza para compararlo con el valor $F_{calculado}$.

En el ejemplo desarrollado como F_{cal} (5,12) > F_{tab} (3,6823) se rechaza H_0.

Con los resultados obtenidos en todos los pasos desarrollados anteriormente se construye la tabla ANOVA de un factor intertratamientos o grupos (Tabla 2.2.5).

Tabla 22. Construcción de la tabla de ANOVA de un factor intertratamientos o grupos.

Fuentes de variación	**Suma de cuadrados**	**gl**	**Cuadrados medios**	**Estadístico $F_{calculado}$**	**$F_{tabulado}$**
Tratamientos o grupos	20,13	2	10,07	5,12	3,6823
Errores	23,60	12	1,97		
Total	**43,73**	**14**			

Cuando el procedimiento se desarrolla con el uso de software estadístico la decisión se encuentra relacionada con el p-valor calculado en la prueba estadística (error cometido realmente) y alfa, o sea, el nivel de significación fijado por el investigador (error planificado). Posteriormente, se realiza la prueba post hoc (pruebas de comparaciones múltiples), dentro de las cuales existen varias alternativas, siendo la prueba Tukey la más utilizada.

5. Prueba post hoc

Prueba realizada a posteriori (adicional) que consiste en un procedimiento estadístico que se aplica después de un ANOVA, cuando este indica que al menos una media de algún tratamiento o grupo difiere de otra. El ANOVA indica si existen o no diferencias globales, aunque la prueba post hoc muestra exactamente dónde se encuentran esas diferencias o similitudes (cuando se obtiene un p-valor significativo).

Objetivo

Identificar de forma específica entre qué pares de tratamientos o grupos existen diferencias estadísticas significativas, aunque existe la posibilidad que entre algunos existan similitudes. El ANOVA indica la existencia de diferencias globales, y la prueba post hoc muestra específicamente donde se encuentran esas diferencias o similitudes.

Cuando se utiliza

1. Después de encontrar diferencias estadísticas significativas en el ANOVA (se obtiene un ANOVA significativo). Se comparan todos los pares de tratamientos o grupos para detectar dónde se

encuentran diferencias o similitudes estadísticas. El ANOVA solamente indica si se presentan o no diferencias estadísticas, sin embargo, cuando detecta diferencias, no especifica dónde se encuentra.

2. En estudios donde se cuenta con más de dos tratamientos o grupos y se necesita realizar comparaciones múltiples. Si existen solamente dos tratamientos o grupos, es suficiente realizar la prueba t de Student.
3. Cuando se busca precisión ya que posibilita identificar patrones más detallados en los datos.

Importancia

1. Evitan conclusiones erróneas al controlar el riesgo de cometer errores Tipo I (falsos positivos) en las comparaciones múltiples. Si no se realiza la prueba Post hoc aumenta el riesgo de encontrar diferencias falsas aumenta al hacer múltiples comparaciones.
2. Se debe realizar la selección de forma adecuada, ya que no tienen el mismo algoritmo; algunas son más conservadoras (Scheffé), otras más sensibles (Tukey).
3. Garantizan el balance entre rigor y poder estadístico, debido a que una prueba muy conservadora reduce la probabilidad de detectar diferencias reales.

Tipos de pruebas post hoc más comunes

En la tabla 2.2.6 se muestran las pruebas post hoc disponibles en el software estadístico Jamovi, sus características y su uso específico.

Tabla 23.Pruebas post hoc implementadas en Jamovi.

Prueba post hoc	Características	Uso específico
Tukey HSD	Controla el error conjunto (Tipo I) al comparar todas las medias posibles por pares.	Común en estudios con tratamientos o grupos balanceados.
Bonferroni	Ajusta el nivel de significancia al dividir alfa (α) entre el número de comparaciones.	Es útil cuando se presentan pocas comparaciones.
Scheffe	Más conservadora, adecuada para muchas comparaciones.	Se utiliza cuando se requiere probar contrastes más generales.
Holm-Bonferroni	Controla el error Tipo I de forma más eficiente que el Bonferroni clásico. Mantiene un buen equilibrio entre rigor y poder estadístico.	Menos conservadora que Bonferroni. Tiene mayor poder para detectar diferencias reales sin aumentar demasiado el riesgo de error. Es útil cuando se presentan muchas comparaciones.

Tamaño del efecto (η^2)

El tamaño del efecto eta cuadrado (η^2), en un ANOVA de un factor intertratamientos, es el valor que permite medir cuanta varianza en la variable dependiente cuantitativa es resultado de la influencia de la variable cualitativa independiente. El η^2 (effect size en inglés) mide la proporción de la varianza total de los datos que es explicada por el factor

de estudio; o sea, mide cuanto afecta la variable independiente (factor de estudio) a la variable dependiente (variable respuesta).

El propósito del η^2 es indicar qué proporción de la variabilidad total en los datos puede atribuirse al factor de estudio (tratamientos o grupos) que se está evaluando. Es una métrica que complementa al p-valor, al mostrar la relevancia práctica de los resultados.

En el ANOVA de un factor intertratamientos o grupos el η^2se calcula mediante la fórmula:

$$\eta^2 = \frac{Suma\ de\ cuadrados\ entre\ tratamientos}{Suma\ de\ cuadrados\ total}$$

Características

- Indica qué proporción de la variabilidad en los datos se explica por el efecto del factor de estudio.
- Depende del p-valor y de la magnitud práctica de la diferencia o relación.
- Una diferencia estadísticamente significativa puede tener un efecto pequeño, y viceversa.

Clasificación del tamaño del efecto

El tamaño del efecto se clasifica según los valores que se obtengan de η^2 y d de Cohen (Tabla 24).

Tabla 24. Niveles de clasificación más empleados para el tamaño del efecto.

Valor de η^2	Valor de d	Categorías	Interpretación
≈ **0,01**	0,20	Pequeño	El factor explica muy poca variabilidad. Efecto despreciable.
≈ **0,06**	0,50	Intermedio	El factor explica una parte moderada de la variabilidad. Efecto relevante.
≈ **0,14 o más**	0,80 o más	Grande	El factor explica gran parte de la variabilidad. Efecto muy fuerte.

Fuente: Cohen (1988).

Importancia

- Complementa el p-valor debido a que muestra la magnitud práctica del efecto que provoca el factor de estudio.
- Permite comparar resultados entre diferentes estudios.
- Indica si las diferencias encontradas son relevantes en la práctica.

Conclusiones

El procedimiento de prueba de hipótesis se desarrolla cuando el enunciado del estudio puede ser calificado de verdadero o falso, o sea, cuando es posible emitir un juicio de valor.

La PH es un procedimiento estadístico basado en la evidencia muestral y en la teoría de la probabilidad empleado para determinar si la suposición o idea es verdadera o falsa. En el planteamiento de hipótesis se plantean dos hipótesis mutuamente excluyentes (H_1 y H_0).

Plantear una hipótesis de investigación que coincida con la hipótesis nula supone una aplicación incorrecta del razonamiento estadístico.

La prueba de hipótesis se desarrolla con la aplicación de cinco pasos, además de la conclusión e interpretación de resultados. Rechazar la hipótesis nula en una comparación de grupos significa que el FE influye de forma significativa en los resultados de la variable estudiada.

Seleccionar adecuadamente entre pruebas paramétricas y no paramétricas asegura que los resultados reflejen la realidad de los datos y respalden decisiones científicas sólidas. La clave está en evaluar los supuestos de los datos antes de elegir la prueba, equilibrando potencia estadística y robustez.

El hallazgo de un p-valor estadísticamente significativo debe interpretarse con cautela. Es necesario complementar este resultado con un análisis más profundo de los datos mediante otras herramientas estadísticas. Evaluar el p-valor de manera aislada suele estar sobrevalorado; por ello, en investigaciones científicas resulta recomendable incluir el cálculo del tamaño del efecto, que aporta información sobre la magnitud y relevancia práctica de los hallazgos.

La prueba post hoc es esencial en estadística porque permite localizar las diferencias específicas entre tratamientos o grupos, después de un ANOVA significativo; garantiza resultados más confiables y evitando errores en la interpretación. En estudios con múltiples grupos, elegir la prueba adecuada (Tukey, Bonferroni, Scheffé, etc.) es clave para obtener conclusiones válidas.

2.4 Práctica 4. Pruebas paramétricas en investigación experimental.

2.4.1 Contenidos

Verificación de supuestos de los modelos paramétricos. Prueba t de Student para grupos independientes. ANOVA de un factor intertratamientos o grupos. Pruebas post hoc. Representación de resultados obtenidos.

2.4.2 Objetivo de aprendizaje

Aplicar correctamente las pruebas paramétricas utilizadas en contextos de investigación experimental mediante la verificación de supuestos del modelo, ejecución de procedimientos adecuados y presentación de resultados de manera técnica y comprensible.

2.4.3 Introducción

En la investigación experimental las pruebas paramétricas constituyen un conjunto fundamental de herramientas estadísticas que permiten evaluar diferencias entre tratamientos o grupos y poner a prueba hipótesis derivadas de modelos teóricos. Para aplicar estas pruebas de manera correcta, es indispensable verificar previamente los supuestos o requisitos que las sustentan (normalidad de los datos, homogeneidad de varianzas e independencia de observaciones). El análisis adecuado de estos supuestos garantiza la validez de las conclusiones obtenidas y evita interpretaciones erróneas de los datos.

En esta práctica se abordarán los procedimientos esenciales para el uso de pruebas paramétricas en contextos experimentales. Se iniciará con la

revisión de los supuestos que deben cumplir los modelos (normalidad de datos y homogeneidad de varianzas), para luego aplicar la prueba t de Student para dos tratamientos o grupos independientes. Posteriormente, se desarrollará el análisis de varianza de un factor intertratamientos o grupos, empleado cuando se requiere comparar más de dos tratamientos o condiciones. Además, se introducirán las pruebas post hoc, necesarias cuando el ANOVA indica la existencia de diferencias significativas, con el fin de identificar específicamente entre qué grupos se presentan dichas diferencias.

Finalmente, se trabajará en la representación adecuada de los resultados mediante tablas y gráficos, a partir de criterios de presentación científica. Se busca fortalecer la comprensión conceptual y el dominio técnico de las pruebas paramétricas, que posibilite promover un procedimiento estadístico riguroso y una interpretación fundamentada de los hallazgos experimentales.

2.4.4 Ejercicio 6. Prueba t de Student para tratamientos o grupos independientes.

Contexto

Se quiere comparar el efecto de dos virus preparados sobre hojas apicales de plantas de tabaco. Para ello, se realizó el siguiente experimento: se seleccionaron al azar ocho plantas de tabaco y en cada una de ellas se tomaron dos hojas apicales, sobre cada una de ellas se aplicaron los dos preparados antivirales, conteniendo los virus cuyos efectos se querían evaluar (ocho hojas con el preparado antiviral 1 y ocho hojas con el preparado antiviral 2). La variable de respuesta medida fue la superficie de las lesiones locales (mm^2) que aparecen como

pequeñas manchas oscuras en las hojas de la planta de tabaco. Utilice una confiabilidad en la estimación del 95 %.

Datos recolectados

Tabla 25.Prueba t de Student para tratamientos o grupos independientes.

Preparados antivirales	Superficie de lesiones locales en la hoja de tabaco (mm^2)							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Preparado 1	31	20	18	17	29	28	10	17
Preparado 2	18	17	14	11	10	7	5	6

Requerimientos del ejercicio

1. Construya la matriz de datos con el uso de Jamovi y desarrollo el procedimiento en el visor de resultados del software estadístico.
2. ¿Los datos siguen una distribución similar a la distribución normal? Concluya e interprete.
3. ¿Se presenta homogeneidad de varianzas en los tratamientos? Concluya e interprete.
4. ¿Cuál es el efecto de los preparados virales sobre hojas apicales de plantas de tabaco?
5. Represente los resultados obtenidos mediante el gráfico correspondiente y establezca letras para definir similitudes o diferencias estadísticas entre los preparados antivirales.

Procedimiento estadístico

1. Descripción de los pasos en Jamovi para construir la matriz de datos y desarrollar el procedimiento

- Abrir el software estadístico Jamovi.

- Renombrar el archivo (ejercicio 4.1) y ubicarlo en la carpeta creada en su ordenador para ubicar los ejercicios.
- Haga clic en el menú variables; y seleccione la opción editar para colocar el nombre de cada variable, unidad de medición, tipo de medida y tipo de dato.
- Haga clic en el menú datos y digite de forma manual los datos correspondientes a las variables preparados antivirales (nominal) y superficie de lesiones en hojas apicales del tabaco que es de escala (mm^2). Las columnas representan a las variables y las filas a las unidades de estudio o de muestreo (casos).
- Haga clic en Pruebas t y seleccione prueba t para muestras independientes (tratamientos o grupos independientes). Pasar a la derecha la variable dependiente (superficie de lesiones) y la variable de agrupación (preparados antivirales) y se seleccionan dentro de pruebas (t de Student), intervalo de confianza (95 %), hipótesis bilateral (Grupo 1 ≠ Grupo 2), descriptivas (opción que entrega los estadísticos descriptivos de cada tratamiento) y la comprobación de supuestos (prueba de homogeneidad de varianzas y prueba de normalidad de datos).

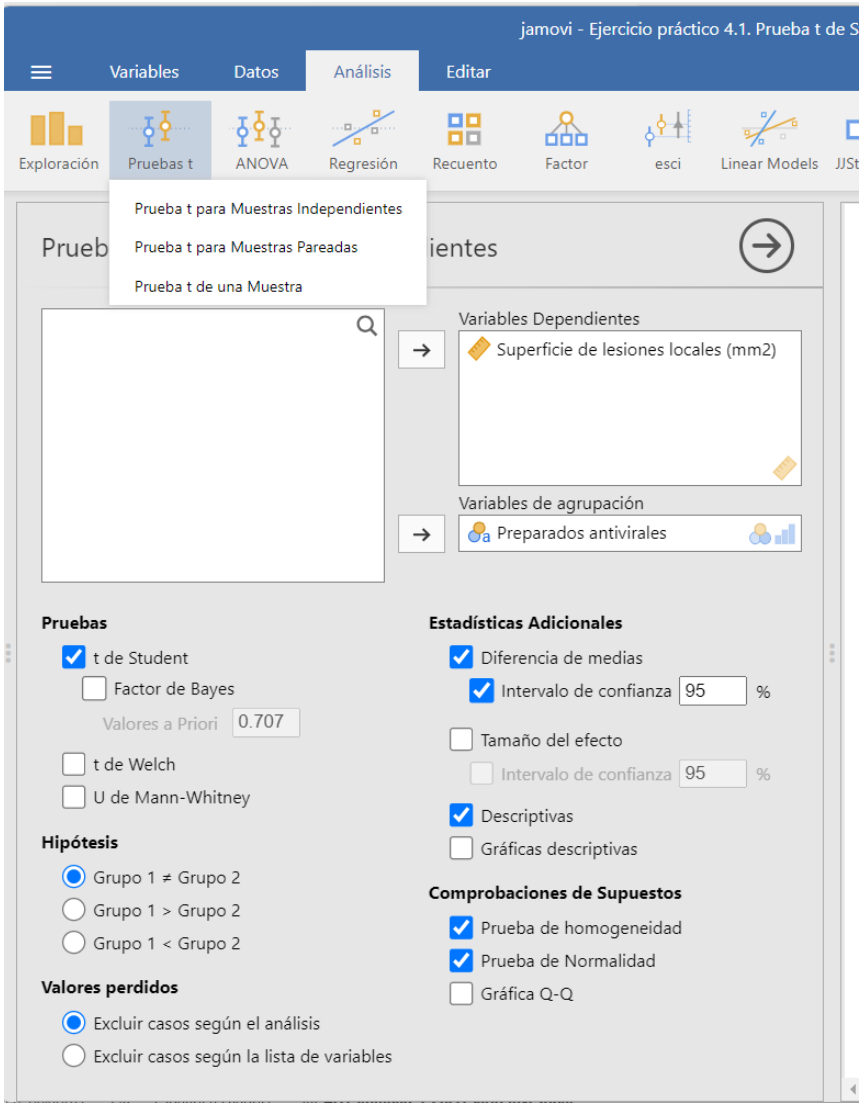


Figura 88. Pruebas t.

2. Verificación del supuesto de normalidad de datos

1. Planteamiento de hipótesis

H_0: Los datos de superficie de las lesiones locales en la hoja de la planta de tabaco siguen una distribución; similar a la distribución normal.

H_1. Los datos de superficie de las lesiones locales en la hoja de la planta de tabaco no siguen una distribución; similar a la distribución normal.

2. Establecimiento del nivel de significancia.

α=0,05. NS=5,0 %.

3. Selección y ejecución de la prueba estadística

Prueba de Shapiro-Wilk.

Tabla 26. Prueba de Shapiro-Wilk.

Variable	W	p-valor
Superficie de lesiones locales en la hoja de tabaco (mm^2)	0,94	0,348

Nota. Un p-valor bajo sugiere una violación del supuesto de normalidad de datos.

4. Lectura del p-valor

El p-valor obtenido en la prueba estadística realizada (0,348) es mayor a alfa (0,05).

5. Tomar la decisión

Dado que el p-valor obtenido es mayor a alfa no se rechaza H_0.

6. Conclusión

Con una confiabilidad en la estimación del 95 % y un error de probabilidad del 34,8 % se concluye que los datos de superficie de las

lesiones locales en la hoja de la planta de tabaco siguen una distribución; similar a la distribución normal.

7. Interpretación

El cumplimiento del supuesto de normalidad de datos garantiza la utilización de la prueba paramétrica t de Student para tratamientos o grupos independientes.

3. Verificación del supuesto de homogeneidad de varianzas

1. Planteamiento de hipótesis

H_0: Las varianzas de los tratamientos antivirales son homogéneas.

H_1. Las varianzas de los tratamientos antivirales no son homogéneas.

2. Establecimiento del nivel de significancia.

α=0,05. NS=5,0 %.

3. Selección y ejecución de la prueba estadística

Prueba de Levene.

Tabla 27.Prueba de Levene.

Variable	F	gl_1	gl_2	p-valor
Superficie de lesiones locales en la hoja de tabaco (mm^2)	1,90	1	14	0,190

Nota. Un p-valor bajo sugiere una violación del supuesto de varianzas iguales.

4. Lectura del p-valor

El p-valor obtenido en la prueba estadística realizada (0,190) es mayor a alfa (0,05).

5. Tomar la decisión

Dado que el p-valor obtenido es mayor a alfa no se rechaza H_0.

6. Conclusión

Con una confiabilidad en la estimación del 95 % y un error de probabilidad del 19,0 % se concluye que las varianzas en los tratamientos antivirales son homogéneas.

7. Interpretación

El cumplimiento del supuesto de homogeneidad de varianzas garantiza la utilización de la prueba paramétrica t de Student para tratamientos o grupos independientes.

4. Comprobación del efecto de los preparados antivirales sobre la superficie de lesiones locales en las hojas apicales de plantas de tabaco

1. Planteamiento de hipótesis

H_0: Las medias de superficie de lesiones locales en las hojas apicales de la planta de tabaco bajo el efecto de preparados antivirales son iguales.

H_1. Las medias de superficie de lesiones locales en las hojas apicales de la planta de tabaco bajo el efecto de preparados antivirales son diferentes.

2. Establecimiento del nivel de significancia.

α=0,05. NS=5,0 %.

3. Selección y ejecución de la prueba estadística.

Prueba t de Student para tratamientos o grupos independientes.

Tabla 28.Prueba t de Student para tratamientos o grupos independientes.

Variable	Estadístico	gl	p-valor	Diferencia de medias	ES de la diferencia
Superficie de lesiones locales en la hoja de tabaco (mm^2)	3,28	14	0,0055	10,25	3,13

ES=Error típico o estándar.

Estadísticos descriptivos de cada tratamiento

Tabla 29.Estadísticos descriptivos de cada tratamiento.

Variable	Tratamientos	n	Media	Mediana	DE	$ES_{\bar{X}}$
Superficie de lesiones locales en la hoja de tabaco (mm^2)	Preparado 1	8	21,25	19,00	7,32	2,59
	Preparado 2	8	11,00	10,50	4,96	1,75

DE=Desviación típica o estándar de los datos. $ES_{\bar{X}}$= Error típico o estándar de la media aritmética.

4. Lectura del p-valor

El p-valor obtenido en la prueba estadística realizada (0,005) es menor a alfa (0,05).

5. Tomar la decisión

Dado que el p-valor obtenido en la prueba estadística es menor a alfa se rechaza H_0; o sea, se acepta la hipótesis del investigador (H_1).

6. Conclusión

Con una confiabilidad en la estimación del 95 % y un error de probabilidad del 0,5 % se concluye que las medias de superficie de lesiones locales en las hojas apicales de la planta de tabaco bajo el efecto de preparados antivirales son diferentes.

7. Interpretación

Las hojas apicales de las plantas de tabaco tratadas con los diferentes preparados mostraron diferencias estadísticas significativas en relación con la superficie de las lesiones locales. El preparado 2 mostró un control superior en las hojas apicales ($\bar{X}$=11,0 mm^2 de superficie de lesiones locales ±4,95 mm^2 y un $EE_{\bar{X}}$=1,75 mm^2) significativamente mayor que el preparado 1 ($\bar{X}$=21,25 mm^2 ±7,32 mm^2 y un $EE_{\bar{X}}$=2,59 mm^2, con un t=3,28 y 14 gl; un p-valor=0,005 (0,5 % de error que se comete) menor al nivel de significancia (α=0,05).

5. Cálculo del tamaño del efecto

Debido a que la prueba t de Student para grupos independientes no informa sobre la magnitud de la diferencia que pueda existir entre los tratamientos se calcula la d de Cohen.

$$d = \frac{X_1 - X_2}{S_p}$$

Cálculo de la desviación estándar combinada (*Sp*), o sea, considerando ambos tratamientos.

$$S_p = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_1 - 1)S_2^2}{n_1 + n_2 - 2}}$$

$$S_p = \sqrt{\frac{7(7{,}324)^2 + 7(4{,}957)^2}{8 + 8 - 2}}$$

$$S_p = \sqrt{\frac{7\,(53{,}64) + 7\,(24{,}57)}{14}}$$

$$S_p = \sqrt{\frac{375{,}48 + 171{,}99}{14}}$$

$$S_p = \sqrt{\frac{547{,}47}{14}}$$

$$S_p = \sqrt{39{,}10}$$

$$S_p = 6{,}25$$

Sustituyendo se obtiene

$$d = \frac{21{,}25 - 11{,}0}{6{,}25}$$

$$d = \frac{10{,}25}{6{,}25}$$

$$d = 1{,}64$$

La diferencia observada entre los tratamientos es aproximadamente 1,64 desviaciones estándar, lo que indica una separación importante entre ambos preparados. Considerando los criterios de Cohen (Tabla 2.2.7) el tamaño del efecto es muy grande (d de Cohen = 1,64), indicando una diferencia sustancial en la respuesta biológica evaluada.

6. Representación gráfica de los resultados obtenidos

En el gráfico se muestran las medias aritméticas de cada preparado antiviral y las letras indican que se presentan diferencias estadísticas significativas (Figura 89).}

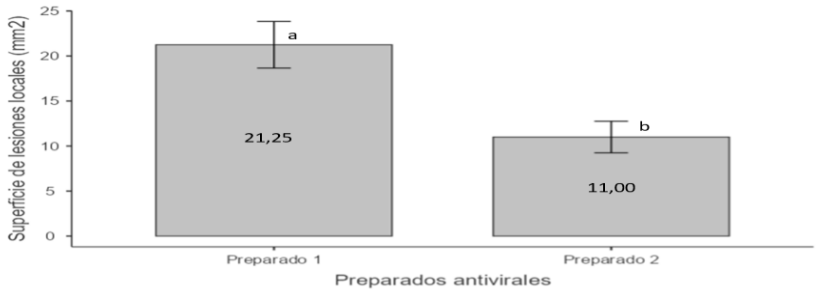


Figura 89.Efecto de preparados antivirales en la superficie de lesiones locales en hojas de tabaco.

2.4.5 Ejercicio 7. ANOVA de un factor intertratamientos o grupos

Contexto

Una empresa agrícola experimenta cuatro fórmulas de fertilizantes para la nutrición del suelo en un área sembrada con el cultivo de tomate; y se necesita establecer si se presenta o no diferencia estadística significativa entre los tratamientos (fórmula aplicada) en función del rendimiento agrícola (t ha^{-1}) del cultivo (Utilice un alfa=0,04).

Tabla 30. Datos recolectados.

Fórmulas de fertilizante	**Observaciones**							
	Rendimiento agrícola (t ha^{-1})							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
F-1	15,3	15,0	16,0	15,2	15,8	15,0	14,8	15,1
F-2	14,8	15,3	15,7	16,5	14,5	14,0	14,7	13,4
F-3	17,5	18,5	18,9	17,0	18,0	19,0	18,2	17,5
F-4	18,7	17,9	19,4	19,7	17,9	20,1	19,7	20,5

Requerimientos del ejercicio

1. Construya la matriz de datos con el uso de Jamovi y desarrolle el procedimiento en el visor de resultados del software estadístico.
2. Verifique los supuestos o requisitos del modelo paramétrico.
3. Determine si se presenta efecto de la fórmula de fertilizante en el rendimiento agrícola del cultivo de tomate.
4. Realice una descripción estadística del rendimiento agrícola para cada tratamiento del factor de estudio.
5. Realice la prueba post hoc para conocer entre que fórmula de fertilizante se presentan diferencias y entre cual similitudes.

6. Represente los resultados obtenidos mediante el gráfico correspondiente y establezca letras para definir similitudes o diferencias estadísticas entre las fórmulas de fertilizantes.

Procedimiento estadístico

1. Descripción de los pasos en Jamovi

- Abrir el software estadístico Jamovi.
- Renombrar el archivo (ejercicio 4.2) y ubicarlo en la carpeta creada en su ordenador para ubicar los ejercicios.
- Hacer clic en el menú variables; y en la opción editar hacer clic para ubicar el nombre de cada variable, unidad de medición, tipo de medida y tipo de dato.
- Hacer clic en el menú datos y digitar de forma manual los datos correspondientes a las variables fórmula de fertilizantes (F-1, F-2, F-3 y F-4 que es nominal) y rendimiento agrícola que es de escala ($t\ ha^{-1}$).
- Hacer clic en ANOVA y seleccionar la opción ANOVA de un factor (intertratamientos o grupos), se pasan a la derecha la variable dependiente (rendimiento agrícola) y la variable de agrupación (fórmula de fertilizantes) y en Varianzas se marca Asumir iguales (Fisher), en Estadísticos adicionales se marcan Tabla y Gráficas descriptivas, en comprobación de Supuestos, se marcan las pruebas de homogeneidad de varianzas y de normalidad de datos). En faltan valores se deja la opción que viene por defecto (excluir casos según el análisis). En pruebas Post hoc se marca Tukey (asume varianzas iguales); que se utiliza para

conocer entre que tratamientos se encuentran las diferencias o similitudes. En Estadísticos se marcan mostrar significación y marcar comparaciones significativas.

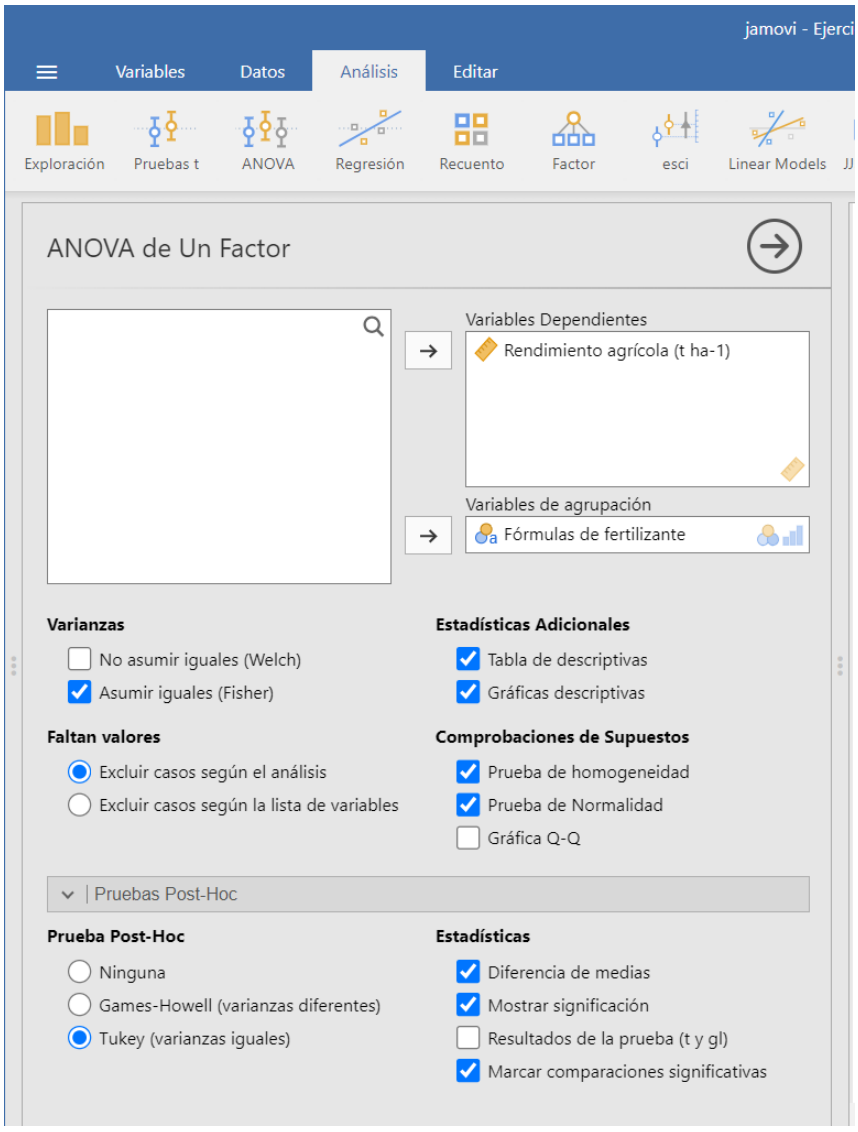


Figura 90.Procedimiento estadístico.

2. Verificación de los supuestos del modelo paramétrico

Normalidad de datos

1. Planteamiento de hipótesis

H_0: Los datos de rendimiento agrícola en el cultivo de tomate siguen una distribución; similar a la distribución normal.

H_1. Los datos de rendimiento agrícola en el cultivo de tomate no siguen una distribución; similar a la distribución normal.

2. Establecimiento del nivel de significancia.

α=0,04. NS=4,0 %. Confiabilidad en la estimación del 96 %.

3. Selección y ejecución de la prueba estadística

Prueba de Shapiro-Wilk.

Tabla 31.Prueba de Shapiro-Wilk.

Variable	W	p-valor
Rendimiento agrícola (t ha^{-1})	0,98	0,859

Nota. Un p-valor bajo sugiere una violación del supuesto de normalidad de datos.

4. Lectura del p-valor

El p-valor obtenido en la prueba estadística realizada (0,859) es mayor a alfa (0,04).

5. Tomar la decisión

Dado que el p-valor obtenido es mayor a alfa no se rechaza H_0.

6. Conclusión

Con una confiabilidad en la estimación del 96 % y un error de probabilidad del 85,9 % se concluye que los datos de rendimiento agrícola en el cultivo de tomate siguen una distribución; similar a la distribución normal.

7. Interpretación

El cumplimiento del supuesto de normalidad de datos garantiza la utilización de la prueba paramétrica ANOVA de un factor intertratamientos o grupos.

Homogeneidad de varianzas

1. Planteamiento de hipótesis

H_0: Las varianzas en cada fórmula de fertilizante son homogéneas.

H_1. Las varianzas en cada fórmula de fertilizante no son homogéneas.

2. Establecimiento del nivel de significancia.

α=0,04. NS=4,0 %.

3. Selección y ejecución de la prueba estadística

Prueba de Levene.

Tabla 32.Prueba de Levene.

Variable	F	gl_1	gl_2	p-valor
Rendimiento agrícola (t ha^{-1})	1,98	3	28	0,141

Nota. Un p-valor bajo sugiere una violación del supuesto de varianzas iguales.

4. Lectura del p-valor

El p-valor obtenido en la prueba estadística realizada (0,141) es mayor a alfa (0,04).

5. Tomar la decisión

Dado que el p-valor obtenido es mayor a alfa no se rechaza H_0.

6. Conclusión

Con una confiabilidad en la estimación del 96 % y un error de probabilidad del 14,1 % se concluye que las varianzas en los tratamientos son homogéneas.

7. Interpretación

El cumplimiento del supuesto de homogeneidad de varianzas garantiza la utilización de la prueba paramétrica ANOVA de un factor intertratamientos o grupos.

3. Comprobación del efecto de la fórmula de fertilizante en el rendimiento agrícola del cultivo de tomate.

1. Planteamiento de hipótesis

H_0: Las medias de rendimiento agrícola del cultivo de tomate bajo el efecto de las fórmulas de fertilizantes son iguales.

H_1: Al menos una media de rendimiento agrícola del cultivo de tomate bajo el efecto de las fórmulas de fertilizantes es diferente.

2. Establecimiento del nivel de significancia.

α=0,04. NS=4,0 %.

3. Selección y ejecución de la prueba estadística.

ANOVA de un factor intertratamientos o grupos (Fisher).

Tabla 33.ANOVA de un factor intertratamientos o grupos (Fisher).

Variable	Estadístico F	gl_1	gl_2	p-valor
Rendimiento agrícola (t ha^{-1})	56,39	3	28	< 0,001

gl_1=grados de libertad inter tratamientos. gl_2= grados de libertad intra tratamientos (errores).

4. Lectura del p-valor

El p-valor obtenido en la prueba estadística realizada (< 0,001) es menor a alfa (0,04).

5. Tomar la decisión

Dado que el p-valor obtenido en la prueba estadística es menor a alfa se rechaza H_0; o sea, se acepta la hipótesis del investigador (H_1).

6. Conclusión

Con una confiabilidad en la estimación del 96 % y un error de probabilidad menor a 0,01 % se concluye que al menos una media de rendimiento agrícola del cultivo de tomate bajo el efecto de las fórmulas de fertilizantes es diferente.

7. Interpretación

La fórmula de fertilizante influye en la variación del rendimiento agrícola del cultivo del tomate.

4. Estadísticos descriptivos de cada tratamiento

Tabla 34. Estadísticos descriptivos de cada tratamiento.

Variable	Fórmulas de fertilizantes	n	Media	DE	$ES_{\bar{X}}$
Rendimiento agrícola (t ha^{-1})	F-1	8	15,28	0,42	0,15
	F-2	8	14,86	0,97	0,34
	F-3	8	18,08	0,71	0,25
	F-4	8	19,24	0,97	0,34

DE=Desviación típica o estándar de los datos. $ES_{\bar{X}}$= Error típico o estándar de la media aritmética.

5. Prueba post hoc

En la prueba post hoc (Tukey) realizada se evidencia que la F-1 ($\bar{X}$=15,3 t ha^{-1}) no presenta diferencias estadísticas significativas con la F-2

($\bar{X}$=14,9 t ha^{-1}); sin embargo, presenta diferencias altamente significativas con la F-3 ($\bar{X}$=18,1 t ha^{-1}) y con la F-4 ($\bar{X}$=19,5 t ha^{-1}). Las F-2, F-3 y F-4 presentan diferencias estadísticas significativas (Tabla 2.2.7).

Tabla 35. Prueba post hoc (Tukey) para fórmulas de fertilizantes y rendimiento agrícola (t ha-1).

Fórmulas de fertilizantes	Diferencia de medias y p-valor	F-1	F-2	F-3	F-4
F-1	Diferencia de medias	—	0,41	-2,80***	-3,96***
	p-valor	—	0,735	<0,001	<0,001
F-2	Diferencia de medias	—	—	-3,21***	-4,37***
	p-valor	—	—	<0,001	<0,001
F-3	Diferencia de medias	—		—	-1,16*
	p-valor	—		—	0,034
F-4	Diferencia de medias	—			—
	p-valor	—			—

Nota. * p < 0,05; ** p < 0,01; *** p < 0,001.

6. Representación gráfica de los resultados

La fórmula de fertilizante 4 (F-4) alcanzó el mayor valor de rendimiento agrícola (19,5 t ha^{-1}) diferente estadísticamente a la F-3 (18,1 t ha^{-1}), F-2 (15, 3 t ha^{-1}) y F-1 (14,9 t ha^{-1}) (Figura 2.2.12).

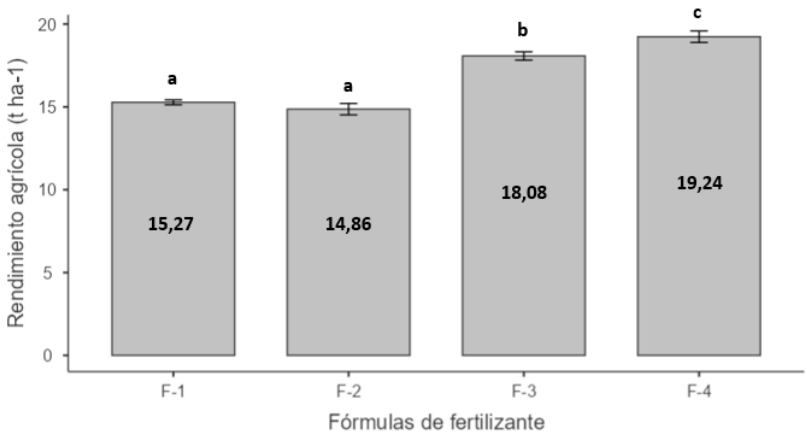


Figura 91.Efecto de la fórmula de fertilizante en el rendimiento agrícola del cultivo de tomate.

*Letras diferentes difieren estadísticamente para un p-valor ≤ 0,04 (Prueba de Tukey).

Conclusiones

La prueba t de Student para tratamientos o grupos independientes es muy utilizada en la práctica, sin embargo, a menudo su aplicación se realiza sin la comprobación de los supuestos que requiere; la falta de normalidad en los datos o de homogeneidad en las varianzas invalida la utilización de la prueba t de Student.

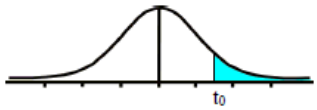
El ANOVA de un factor intertratamientos o grupos es una técnica obligada relacionada con el diseño de experimentos y que se aplica para concluir e interpretar investigaciones cuando el investigador manipula un factor de estudio con más de dos tratamientos y la variable dependiente es numérica y cumple con los requisitos del modelo paramétrico. Cuando se presentan diferencias estadísticas significativas

entre los tratamientos objeto de estudio el ANOVA constituye un proceso intermedio que se debe complementar con pruebas post hoc, las cuales permiten definir entre que tratamientos o grupos se encuentran las diferencias o similitudes.

ANEXOS

Anexo 1. Valores críticos de la Distribución t de Student.

Tabla t-Student



Grados de libertad	0.25	0.1	0.05	0.025	0.01	0.005
1	1.0000	3.0777	6.3137	12.7062	31.8210	63.6559
2	0.8165	1.8856	2.9200	4.3027	6.9645	9.9250
3	0.7649	1.6377	2.3534	3.1824	4.5407	5.8408
4	0.7407	1.5332	2.1318	2.7765	3.7469	4.6041
5	0.7267	1.4759	2.0150	2.5706	3.3649	4.0321
6	0.7176	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.7111	1.4149	1.8946	2.3646	2.9979	3.4995
8	0.7064	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.7027	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.6998	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.6974	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.6955	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.6938	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.6924	1.3450	1.7613	2.1448	2.6245	2.9768
15	0.6912	1.3406	1.7531	2.1315	2.6025	2.9467
16	0.6901	1.3368	1.7459	2.1199	2.5835	2.9208
17	0.6892	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.6884	1.3304	1.7341	2.1009	2.5524	2.8784
19	0.6876	1.3277	1.7291	2.0930	2.5395	2.8609
20	0.6870	1.3253	1.7247	2.0860	2.5280	2.8453
21	0.6864	1.3232	1.7207	2.0796	2.5176	2.8314
22	0.6858	1.3212	1.7171	2.0739	2.5083	2.8188
23	0.6853	1.3195	1.7139	2.0687	2.4999	2.8073
24	0.6848	1.3178	1.7109	2.0639	2.4922	2.7970
25	0.6844	1.3163	1.7081	2.0595	2.4851	2.7874
26	0.6840	1.3150	1.7056	2.0555	2.4786	2.7787
27	0.6837	1.3137	1.7033	2.0518	2.4727	2.7707
28	0.6834	1.3125	1.7011	2.0484	2.4671	2.7633
29	0.6830	1.3114	1.6991	2.0452	2.4620	2.7564
30	0.6828	1.3104	1.6973	2.0423	2.4573	2.7500
31	0.6825	1.3095	1.6955	2.0395	2.4528	2.7440
32	0.6822	1.3086	1.6939	2.0369	2.4487	2.7385
33	0.6820	1.3077	1.6924	2.0345	2.4448	2.7333
34	0.6818	1.3070	1.6909	2.0322	2.4411	2.7284
35	0.6816	1.3062	1.6896	2.0301	2.4377	2.7238
36	0.6814	1.3055	1.6883	2.0281	2.4345	2.7195
37	0.6812	1.3049	1.6871	2.0262	2.4314	2.7154
38	0.6810	1.3042	1.6860	2.0244	2.4286	2.7116
39	0.6808	1.3036	1.6849	2.0227	2.4258	2.7079
40	0.6807	1.3031	1.6839	2.0211	2.4233	2.7045
41	0.6805	1.3025	1.6829	2.0195	2.4208	2.7012
42	0.6804	1.3020	1.6820	2.0181	2.4185	2.6981
43	0.6802	1.3016	1.6811	2.0167	2.4163	2.6951
44	0.6801	1.3011	1.6802	2.0154	2.4141	2.6923
45	0.6800	1.3007	1.6794	2.0141	2.4121	2.6896
46	0.6799	1.3002	1.6787	2.0129	2.4102	2.6870
47	0.6797	1.2998	1.6779	2.0117	2.4083	2.6846
48	0.6796	1.2994	1.6772	2.0106	2.4066	2.6822
49	0.6795	1.2991	1.6766	2.0096	2.4049	2.6800

50	0.6794	1.2987	1.6759	2.0086	2.4033	2.6778
51	0.6793	1.2984	1.6753	2.0076	2.4017	2.6757
52	0.6792	1.2980	1.6747	2.0066	2.4002	2.6737
53	0.6791	1.2977	1.6741	2.0057	2.3988	2.6718
54	0.6791	1.2974	1.6736	2.0049	2.3974	2.6700
55	0.6790	1.2971	1.6730	2.0040	2.3961	2.6682
56	0.6789	1.2969	1.6725	2.0032	2.3948	2.6665
57	0.6788	1.2966	1.6720	2.0025	2.3936	2.6649
58	0.6787	1.2963	1.6716	2.0017	2.3924	2.6633
59	0.6787	1.2961	1.6711	2.0010	2.3912	2.6618
60	0.6786	1.2958	1.6706	2.0003	2.3901	2.6603
61	0.6785	1.2956	1.6702	1.9996	2.3890	2.6589
62	0.6785	1.2954	1.6698	1.9990	2.3880	2.6575
63	0.6784	1.2951	1.6694	1.9983	2.3870	2.6561
64	0.6783	1.2949	1.6690	1.9977	2.3860	2.6549
65	0.6783	1.2947	1.6686	1.9971	2.3851	2.6536
66	0.6782	1.2945	1.6683	1.9966	2.3842	2.6524
67	0.6782	1.2943	1.6679	1.9960	2.3833	2.6512
68	0.6781	1.2941	1.6676	1.9955	2.3824	2.6501
69	0.6781	1.2939	1.6672	1.9949	2.3816	2.6490
70	0.6780	1.2938	1.6669	1.9944	2.3808	2.6479
71	0.6780	1.2936	1.6666	1.9939	2.3800	2.6469
72	0.6779	1.2934	1.6663	1.9935	2.3793	2.6458
73	0.6779	1.2933	1.6660	1.9930	2.3785	2.6449
74	0.6778	1.2931	1.6657	1.9925	2.3778	2.6439
75	0.6778	1.2929	1.6654	1.9921	2.3771	2.6430
76	0.6777	1.2928	1.6652	1.9917	2.3764	2.6421
77	0.6777	1.2926	1.6649	1.9913	2.3758	2.6412
78	0.6776	1.2925	1.6646	1.9908	2.3751	2.6403
79	0.6776	1.2924	1.6644	1.9905	2.3745	2.6395
80	0.6776	1.2922	1.6641	1.9901	2.3739	2.6387
81	0.6775	1.2921	1.6639	1.9897	2.3733	2.6379
82	0.6775	1.2920	1.6636	1.9893	2.3727	2.6371
83	0.6775	1.2918	1.6634	1.9890	2.3721	2.6364
84	0.6774	1.2917	1.6632	1.9886	2.3716	2.6356
85	0.6774	1.2916	1.6630	1.9883	2.3710	2.6349
86	0.6774	1.2915	1.6628	1.9879	2.3705	2.6342
87	0.6773	1.2914	1.6626	1.9876	2.3700	2.6335
88	0.6773	1.2912	1.6624	1.9873	2.3695	2.6329
89	0.6773	1.2911	1.6622	1.9870	2.3690	2.6322
90	0.6772	1.2910	1.6620	1.9867	2.3685	2.6316
91	0.6772	1.2909	1.6618	1.9864	2.3680	2.6309
92	0.6772	1.2908	1.6616	1.9861	2.3676	2.6303
93	0.6771	1.2907	1.6614	1.9858	2.3671	2.6297
94	0.6771	1.2906	1.6612	1.9855	2.3667	2.6291
95	0.6771	1.2905	1.6611	1.9852	2.3662	2.6286
96	0.6771	1.2904	1.6609	1.9850	2.3658	2.6280
97	0.6770	1.2903	1.6607	1.9847	2.3654	2.6275
98	0.6770	1.2903	1.6606	1.9845	2.3650	2.6269
99	0.6770	1.2902	1.6604	1.9842	2.3646	2.6264
100	0.6770	1.2901	1.6602	1.9840	2.3642	2.6259
∞	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758

Figura 92. Valores críticos de la Distribución t de Student.

Anexo 2. Valores críticos de la distribución F para los niveles de significación más frecuentes.

Valores críticos de la distribución F para P=0,05

DF ←Denominador	Numerador DF												
	1	2	3	4	5	7	10	15	20	30	60	120	500
1	161.45	199,50	215.71	224.58	230.16	236.77	241.88	245,95	248.01	250.10	252.20	253.25	254.06
2	18.513	19.000	19.164	19.247	19.296	19.353	19.396	19.429	19.446	19.462	19.479	19.487	19.494
3	10.128	9.5522	9.2766	9.1172	9.0135	8.8867	8.7855	8.7028	8.6602	8.6165	8.5720	8.5493	8.5320
4	7.7086	6.9443	6.5915	6.3882	6.2560	6.0942	5.9644	5.8579	5.8026	5.7458	5.6877	5.6580	5.6352
5	6.6078	5.7862	5.4095	5.1922	5.0504	4.8759	4.7351	4.6187	4.5582	4.4958	4.4314	4.3985	4.3731
7	5.5914	4.7375	4.3469	4.1202	3.9715	3.7871	3.6366	3.5108	3.4445	3.3758	3.3043	3.2675	3.2388
10	4.9645	4.1028	3.7082	3.4780	3.3259	3.1354	2.9782	2.8450	2.7741	2.6996	2.6210	2.5801	2.5482
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7066	2.5437	2.4035	2.3275	2.2467	2.1601	2.1141	2.0776
20	4.3512	3.4928	3.0983	2.8660	2.7109	2.5140	2.3479	2.2032	2.1241	2.0391	1.9463	1.8962	1.8563
30	4.1709	3.3159	2.9223	2.6896	2.5336	2.3343	2.1646	2.0149	1.9317	1.8408	1.7396	1.6835	1.6376
60	4.0012	3.1505	2.7581	2.5252	2.3683	2.1666	1.9927	1.8365	1.7480	1.6492	1.5343	1.4672	1.4093
120	3.9201	3.0718	2.6802	2.4473	2.2898	2.0868	1.9104	1.7505	1.6587	1.5544	1.4289	1.3519	1.2804
500	3.8601	3.0137	2.6227	2.3898	2.2320	2.0278	1.8496	1.6864	1.5917	1.4820	1.3455	1.2552	1.1586

Valores críticos de la distribución F para P=0,10

DF ←Denominador	Numerador DF												
	1	2	3	4	5	7	10	15	20	30	60	120	500
1	39.864	49.500	53.593	55.833	57.240	58.906	60.195	61.220	61.740	62.265	62.794	63.061	63.264
2	8.5264	8.9999	9.1618	9.2434	9.2926	9.3491	9.3915	9.4248	9.4413	9.4580	9.4745	9.4829	9.4893
3	5.5384	5.4624	5.3907	5.3426	5.3092	5.2661	5.2304	5.2003	5.1845	5.1681	5.1513	5.1425	5.1358
4	4.5448	4.3245	4.1909	4.1073	4.0505	3.9790	3.9198	3.8704	3.8443	3.8175	3.7896	3.7753	3.7643
5	4.0605	3.7798	3.6194	3.5202	3.4530	3.3679	3.2974	3.2379	3.2067	3.1740	3.1402	3.1228	3.1094
7	3.5895	3.2575	3.0740	2.9605	2.8833	2.7850	2.7025	2.6322	2.5947	2.5555	2.5142	2.4927	2.4761
10	3.2850	2.9244	2.7277	2.6054	2.5216	2.4139	2.3226	2.2434	2.2007	2.1554	2.1071	2.0818	2.0618
15	3.0731	2.6951	2.4898	2.3615	2.2729	2.1582	2.0593	1.9722	1.9243	1.8727	1.8168	1.7867	1.7629
20	2.9746	2.5893	2.3801	2.2490	2.1582	2.0397	1.9368	1.8450	1.7939	1.7383	1.6768	1.6432	1.6163
30	2.8808	2.4887	2.2761	2.1423	2.0493	1.9269	1.8195	1.7222	1.6674	1.6064	1.5376	1.4990	1.4669
60	2.7911	2.3932	2.1774	2.0409	1.9457	1.8194	1.7070	1.6034	1.5435	1.4756	1.3953	1.3476	1.3060
120	2.7478	2.3473	2.1300	1.9924	1.8959	1.7675	1.6523	1.5450	1.4821	1.4094	1.3203	1.2646	1.2123
500	2.7157	2.3132	2.0947	1.9561	1.8588	1.7288	1.6115	1.5009	1.4354	1.3583	1.2600	1.1937	1.1215

Valores críticos de la distribución F para P=0,02

DF ↓ Denominador / Numerador DF	1	2	3	4	5	7	10	15	20	30	60	120	500
1	1012.5	1249.5	1350.5	1405.8	1440.6	1481.8	1513.7	1539.1	1551.9	1564.9	1578.0	1584.6	1589.6
2	48.505	49.000	49.166	49.249	49.299	49.356	49.398	49.432	49.448	49.465	49.482	49.490	49.496
3	20.618	18.858	18.110	17.694	17.429	17.110	16.860	16.657	16.553	16.448	16.340	16.286	16.244
4	14.040	12.142	11.344	10.899	10.616	10.274	10.003	9.7828	9.6696	9.5540	9.4359	9.3760	9.3300
5	11.323	9.4544	8.6702	8.2330	7.9530	7.6137	7.3438	7.1234	7.0094	6.8926	6.7728	6.7119	6.6649
7	8.9877	7.2026	6.4539	6.0347	5.7647	5.4354	5.1711	4.9531	4.8392	4.7220	4.6007	4.5384	4.4902
10	7.6384	5.9336	5.2182	4.8157	4.5549	4.2347	3.9750	3.7580	3.6437	3.5245	3.3999	3.3354	3.2850
15	6.7730	5.1355	4.4475	4.0584	3.8052	3.4917	3.2345	3.0168	2.9003	2.7775	2.6468	2.5780	2.5237
20	6.3907	4.7875	4.1134	3.7312	3.4817	3.1713	2.9149	2.6955	2.5771	2.4509	2.3148	2.2421	2.1841
30	6.0382	4.4695	3.8093	3.4339	3.1877	2.8803	2.6239	2.4020	2.2805	2.1493	2.0047	1.9255	1.8611
60	5.7127	4.1785	3.5319	3.1633	2.9207	2.6157	2.3586	2.1326	2.0067	1.8676	1.7085	1.6169	1.5383
120	5.5594	4.0423	3.4026	3.0372	2.7963	2.4923	2.2347	2.0059	1.8769	1.7322	1.5613	1.4577	1.3629
500	5.4467	3.9428	3.3083	2.9453	2.7057	2.4024	2.1441	1.9128	1.7809	1.6307	1.4468	1.3273	1.2019

F Distribution critical values for P=0.01

DF ↓ Denominator / Numerator DF	1	2	3	4	5	7	10	15	20	30	60	120	500
1	4052.2	4999.5	5403.4	5624.6	5763.6	5928.4	6055.8	6157.3	6208.7	6260.6	6313.0	6339.4	6359.5
2	98.503	99.000	99.166	99.249	99.299	99.356	99.399	99.433	99.449	99.466	99.482	99.491	99.497
3	34.116	30.817	29.457	28.710	28.237	27.672	27.229	26.872	26.690	26.504	26.316	26.221	26.148
4	21.198	18.000	16.694	15.977	15.522	14.976	14.546	14.198	14.020	13.838	13.652	13.558	13.486
5	16.258	13.274	12.060	11.392	10.967	10.455	10.051	9.7222	9.5526	9.3793	9.2020	9.1118	9.0424
7	12.246	9.5467	8.4513	7.8466	7.4605	6.9929	6.6201	6.3143	6.1554	5.9920	5.8236	5.7373	5.6707
10	10.044	7.5594	6.5523	5.9944	5.6363	5.2001	4.8492	4.5582	4.4055	4.2469	4.0818	3.9964	3.9303
15	8.6831	6.3588	5.4169	4.8932	4.5557	4.1416	3.8049	3.5223	3.3719	3.2141	3.0471	2.9594	2.8906
20	8.0960	5.8489	4.9382	4.4306	4.1027	3.6987	3.3682	3.0880	2.9377	2.7785	2.6078	2.5167	2.4446
30	7.5624	5.3903	4.5098	4.0179	3.6990	3.3046	2.9791	2.7002	2.5486	2.3859	2.2078	2.1108	2.0321
60	7.0771	4.9774	4.1259	3.6491	3.3388	2.9530	2.6318	2.3522	2.1978	2.0284	1.8362	1.7264	1.6328
120	6.8509	4.7865	3.9490	3.4795	3.1736	2.7918	2.4720	2.1914	2.0345	1.8600	1.6557	1.5330	1.4215
500	6.6858	4.6479	3.8210	3.3569	3.0539	2.6751	2.3564	2.0746	1.9152	1.7353	1.5175	1.3774	1.2317

F Distribution critical values for P=0.005

DF ↙Denominator	Numerator DF												
	1	2	3	4	5	7	10	15	20	30	60	120	500
1	16211	19999	21615	22500	23056	23715	24224	24630	24836	25044	25253	25359	25439
2	198.50	199.00	199.17	199.25	199.30	199.36	199.40	199.43	199.45	199.47	199.48	199.49	199.50
3	55.552	49.799	47.467	46.195	45.392	44.434	43.686	43.085	42.777	42.466	42.149	41.989	41.867
4	31.333	26.284	24.259	23.155	22.456	21.622	20.967	20.438	20.167	19.891	19.611	19.468	19.359
5	22.785	18.314	16.530	15.556	14.940	14.200	13.618	13.146	12.903	12.656	12.402	12.274	12.175
7	16.235	12.404	10.882	10.050	9.5221	8.8853	8.3803	7.9677	7.7539	7.5345	7.3088	7.1933	7.1044
10	12.826	9.4270	8.0807	7.3428	6.8723	6.3025	5.8467	5.4706	5.2740	5.0705	4.8592	4.7501	4.6656
15	10.798	7.7007	6.4760	5.8029	5.3721	4.8473	4.4235	4.0697	3.8826	3.6868	3.4802	3.3722	3.2874
20	9.9439	6.9865	5.8176	5.1744	4.7616	4.2569	3.8470	3.5020	3.3178	3.1234	2.9159	2.8058	2.7186
30	9.1796	6.3547	5.2387	4.6233	4.2275	3.7416	3.3439	3.0058	2.8231	2.6277	2.4151	2.2998	2.2066
60	8.4946	5.7950	4.7290	4.1399	3.7599	3.2911	2.9042	2.5705	2.3872	2.1874	1.9621	1.8341	1.7256
120	8.1789	5.5393	4.4972	3.9207	3.5482	3.0874	2.7052	2.3728	2.1882	1.9840	1.7468	1.6055	1.4778
500	7.9498	5.3548	4.3304	3.7632	3.3963	2.9414	2.5625	2.2303	2.0441	1.8352	1.5844	1.4245	1.2595

F Distribution critical values for P=0.001

DF ↙Denominator	Numerator DF												
	1	2	3	4	5	7	10	15	20	30	60	120	500
1	405284	499999	540379	562500	576405	592873	605621	615764	620908	626099	631337	633972	635983
2	998.50	999.00	999.17	999.25	999.30	999.36	999.40	999.43	999.45	999.47	999.48	999.49	999.50
3	167.03	148.50	141.11	137.10	134.58	131.58	129.25	127.37	126.42	125.45	124.47	123.97	123.59
4	74.137	61.245	56.177	53.436	51.712	49.658	48.053	46.761	46.100	45.429	44.746	44.400	44.135
5	47.181	37.122	33.202	31.085	29.752	28.163	26.917	25.911	25.395	24.869	24.333	24.061	23.852
7	29.245	21.689	18.772	17.198	16.206	15.019	14.083	13.324	12.932	12.530	12.119	11.909	11.747
10	21.040	14.905	12.553	11.283	10.481	9.5174	8.7539	8.1288	7.8038	7.4688	7.1224	6.9443	6.8065
15	16.587	11.339	9.3352	8.2526	7.5673	6.7408	6.0808	5.5351	5.2484	4.9502	4.6378	4.4749	4.3478
20	14.819	9.9526	8.0984	7.0960	6.4606	5.6920	5.0753	4.5618	4.2900	4.0051	3.7030	3.5439	3.4184
30	13.293	8.7734	7.0544	6.1245	5.5339	4.8173	4.2389	3.7528	3.4928	3.2171	2.9197	2.7595	2.6310
60	11.973	7.7678	6.1712	5.3067	4.7565	4.0864	3.5415	3.0781	2.8265	2.5549	2.2522	2.0821	1.9390
120	11.380	7.3212	5.7814	4.9471	4.4157	3.7669	3.2372	2.7833	2.5345	2.2621	1.9502	1.7668	1.6027
500	10.957	7.0041	5.5056	4.6935	4.1757	3.5424	3.0234	2.5759	2.3282	2.0538	1.7292	1.5260	1.3191

Figura 93.Valores críticos de la distribución F para los niveles de significación más frecuentes.

BIBLIOGRAFÍAS

Campbell, D. T., & Stanley, J. C. (1995). *Diseños experimentales y cuasiexperimentales en la investigación social* (7.ª ed.). Rand McNally & Company. https://knowledgesociety.usal.es/sites/default/files/campbell-stanley-disec3b1os-experimentales-y-cuasiexperimentales-en-la-investigacic3b3n-social.pdf

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. https://utstat.toronto.edu/~brunner/oldclass/378f16/readings/CohenPower.pdf

Concepto. (2025). *Experimentación científica*. Enciclopedia Concepto. https://concepto.de/experimentacion-cientifica/

Díaz, A. (2009). *Diseño estadístico de experimentos* (2.ª ed.). Universidad de Antioquia. https://books.google.com.ec/books?id=0x0DW6dNiyAC&printsec=frontcover&hl=es

FAO. (2010). *Capítulo 1. Diseño experimental*. Organización de las Naciones Unidas para la Alimentación y la Agricultura. https://www.fao.org/4/t0848s/t0848s03.htm

Fisher, R. A. (1935). *The design of experiments*. Oliver and Boyd. https://books.google.com.ec/books/about/The_Design_of_Experiments.html?id=-EsNAQAAIAAJ

Fisher, R. A. (1950). *Statistical methods for research workers*. https://dn790007.ca.archive.org/0/items/dli.ernet.233797/233797-Statistical%20Methods%20For%20Research%20Workers.pdf

Flores, E. (2012). *Estadística inferencial*. Universidad. https://www.academia.edu/31909638/Estadistica_inferencial

Gastwirth, J. L., Gel, Y. R., & Miao, W. (2009). The impact of Levene's test of equality of variances on statistical theory and practice.

Statistical Science, *24*(3), 343–360. https://doi.org/10.1214/09-STS301

Gómez, K. A., & Gómez, A. A. (1984). *Statistical procedures for agricultural research* (2nd ed.). John Wiley & Sons. https://www.academia.edu/2456341/statistical_Procedures_for_Agricultural_Research

Gordón-Mendoza, R., & Camargo-Buitrago, I. (2015). Selección de estadísticos para la estimación de la precisión experimental en ensayos de maíz. *Agronomía Mesoamericana*, *26*(1), 55–63. https://www.redalyc.org/journal/437/43758353006/html/

Hernández-Sampieri, R., Fernández, C., & Baptista, P. (1991). *Metodología de la investigación* (6.ª ed.). McGraw-Hill. https://apiperiodico.jalisco.gob.mx/api/sites/periodicooficial.jalisco.gob.mx/files/metodologia_de_la_investigacion_-_roberto_hernandez_sampieri.pdf

Hurtado, J. (2000). *Metodología de la investigación holística* (3.ª ed.). Fundación Sypal. https://dariososafoula.files.wordpress.com/2017/01/hurtado-de-barrera-metodologicc81a-de-la-investigaciocc81n-guicc81a-para-la-comprensiocc81n-holicc81stica-de-la-ciencia.pdf

Johnson, R., & Kuby, P. (2012). *Estadística elemental* (11.ª ed.). Cengage Learning. https://www.academia.edu/40310969/Estadistica_Elemental_Jhonson_Kuby

Levene, H. (1960). *Levene test for equality of variances*. National Institute of Standards and Technology. https://www.itl.nist.gov/div898/handbook/eda/section3/eda35a.htm

Lind, D. A., Marchal, W. G., & Wathen, S. A. (2012). *Estadística aplicada a los negocios y la economía* (15.ª ed.). McGraw-Hill. https://gc.scalahed.com/recursos/files/r161r/w25633w/Estadistica%20para%20Administracion%20y%20Negocios.pdf

Manterola, C., & Otzen, T. (2015). Estudios experimentales 2.ª parte. Estudios cuasi-experimentales. *International Journal of Morphology, 33*(1), 382–387. https://www.scielo.cl/scielo.php?pid=S0717-95022015000100060&script=sci_arttext

Montgomery, D. C. (2013). *Design and analysis of experiments* (8th ed.). John Wiley & Sons. https://www.academia.edu/9101936/Diseño_y_análisis_de_experimentos_Douglas_C._Montgomery

OMS. (2026). *IMC clasificación: Tabla completa según la OMS*. https://calcularimc.blog/blog/imc-clasificacion/

Pajankar, V. D. (2022). *Easy statistics with Jamovi: Use of Jamovi for beginners*. SK Print House. https://www.researchgate.net/publication/368387691_Easy_Statistics_with_Jamovi_Use_of_Jamovi_for_beginners

Patel, J. K., Patel, N. M., & Shiyani, R. L. (2001). Coefficient of variation in field experiments and yardstick thereof: An empirical study. *Current Science, 81*(9), 1163–1164. https://www.jstor.org/stable/24106429

Peter, C. R. (2026). *Statistics: Open for everyone*. LibreTexts. https://commons.libretexts.org/book/stats-49315

Pimentel, F. (1990). *Curso de estatística experimental* (13.ª ed.). Livraria Nobel S.A.

Rodríguez, I., Serrano, B., & Villaseñor, D. R. (2018). Estadística básica con datos agropecuarios. En UTMACH (Ed.), *Análisis de datos agropecuarios* (p. 103). Universidad Técnica de Machala. https://repositorio.utmachala.edu.ec/items/f7f6dcd5-4ba0-458b-a146-38354a807623

Ruiz-Ramírez, J. (2009). Relative efficiency and quality of fertilization experiments in sugar cane. *Terra Latinoamericana, 28*, 149–154. https://www.scielo.org.mx/pdf/tl/v28n2/v28n2a6.pdf

Salcedo, A. (2013). *Estadística en la investigación: Competencia transversal en la formación universitaria*. Programa de Cooperación Interfacultades. https://www.researchgate.net/publication/274961332_Estadistica_en_la_Investigacion_competencia_transversal_en_la_formacion_universitaria

Triola, M. F. (2018). *Estadística* (12.ª ed.). Pearson. https://www.academia.edu/44058808/Estadistica_Mario_F_Triola_12ED

Diseño experimental introducción a la experimentación y procesamiento de datos en investigaciones agropecuarias, se publicó en el mes de noviembre de 2026.

ISBN: 978-9907-852-14-1

Grupo Editorial BLR
Ecuador
Cel: +593 98 320 4362
https://grupoblr.com/
publicaciones@grupoblr.com

BIOGRAFÍA DE AUTORES

Irán Rodríguez Delgado:
Profesor titular en la UTMACH. Ingeniero Agrónomo (1992). Máster en Agricultura Sostenible (2009). Ha coordinado y concluido proyectos de investigación en la UTMACH. Se encuentra en la etapa final de sus estudios de doctorado en la Universidad de Matanzas, Cuba.

Hipólito Israel Pérez Iglesias:
Profesor jubilado en la UTMACH. Ingeniero Agrónomo (1976). Doctor en Ciencias Agrícolas (1982). Ha coordinado y concluido proyectos de investigación en la UTMACH.

DISEÑO EXPERIMENTAL INTRODUCCIÓN A LA EXPERIMENTACIÓN Y PROCESAMIENTO DE DATOS EN INVESTIGACIONES AGROPECUARIAS

Estimado lector, el diseño experimental constituye un pilar metodológico esencial en la investigación científica al posibilitar la determinación de relaciones de causalidad mediante la manipulación controlada de variables independientes, la medición precisa de efectos en variables dependientes y la contención del error experimental. A partir del desarrollo del método científico moderno y de las contribuciones estadísticas de Ronald A. Fisher en el siglo XX —en particular los principios de aleatorización, replicación y control local—, el método experimental consolidó una estructura lógica rigurosa que asegura la validez interna y la reproducibilidad de los hallazgos empíricos.

Esta obra surge para responder a las necesidades formativas del syllabus de Diseño Experimental en la Facultad de Ciencias Agropecuarias de la Universidad Técnica de Machala (UTMACH), orientada a estudiantes de Agronomía, Medicina Veterinaria, Acuicultura y Agropecuaria. El texto articula de forma equilibrada los fundamentos conceptuales con la aplicación práctica, detallando etapas críticas como la formulación de hipótesis, la selección de unidades de estudio, el diseño de matrices de datos y el análisis computacional mediante software estadístico especializado.

En su conjunto, el libro aborda modelos experimentales unifactoriales y factoriales aplicados a problemáticas del sector productivo y biológico, posicionando al diseño experimental como una estrategia metodológica integral que fortalece el rigor, la objetividad y la calidad de la evidencia científica en el ámbito agropecuario.

GRUPO BLR

Grupo Editorial BLR
Ecuador
Cel: +593 98 320 4362
https://grupoblr.com/
publicaciones@grupoblr.com

ISBN: 978-9907-852-14-1

UNIVERSIDAD TÉCNICA DE MACHALA
1969