



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO

**MAESTRÍA EN CIENCIAS AGROPECUARIAS Y RECURSOS
NATURALES**

**SUBMUESTREO BALANCEADO EN UN ARREGLO DE PARCELAS
DIVIDIDAS EN CUADRO LATINO PARA UNA SERIE DE
EXPERIMENTOS**

TESIS

**QUE PARA OBTENER EL GRADO DE MAESTRO EN CIENCIAS
AGROPECUARIAS Y RECURSOS NATURALES**

PRESENTA:

JOSÉ ANTONIO RODRÍGUEZ GONZÁLEZ

COMITÉ DE TUTORES:

DR. ANDRÉS GONZÁLEZ HUERTA. Tutor Académico

DRA. DELFINA DE JESÚS PÉREZ LÓPEZ. Tutora Adjunta

M. EN F. ARTEMIO BALBUENA MELGAREJO. Tutor Adjunto

El Cerrillo Piedras blancas, Estado de México, febrero de 2025



PCARN
Programa en Ciencias Agropecuarias y Recursos Naturales

CONTENIDO

	Pág.
AGRADECIMIENTOS.....	iii
DEDICATORIAS	iv
CONTENIDO.....	v
RESUMEN GENERAL	vii
I. INTRODUCCIÓN GENERAL.....	1
II. REVISIÓN DE LITERATURA	4
2.1 Definición de conceptos.....	4
2.1.1 Submuestreo	4
2.1.2 Diseño experimental y factorial	5
2.1.3 Diseño Cuadro o cuadrado latino	8
2.1.4 Parcelas divididas o “split-plot Design”.....	9
2.1.5 Serie de experimentos	11
2.2 Importancia de la investigación	14
2.3 Modelaje	15
2.4 El análisis de varianza.....	17
2.5 Comparación de medias de tratamientos	20
2.5.1 Comparaciones <i>a priori</i>	22

2.5.2 Comparaciones <i>a posteriori</i>	24
2.6 Uso de Software	25
III. MATERIALES Y MÉTODOS	27
3.1 Simbología utilizada	27
3.2 Modelo estadístico	27
3.3 Diseño experimental	28
3.4 Software recomendable.....	28
3.5 Resultados.....	29
IV. RESULTADOS Y DISCUSIÓN.....	30
4.1 Primer articulo	30
4.2 Segundo articulo	50
4.3 Capitulo de libro	71
V. CONCLUSIONES GENERALES	101
VI. BIBLIOGRAFÍA CITADA GENERAL	102

RESUMEN GENERAL

En las últimas décadas se ha preguntado frecuentemente en diversos escenarios la relevancia que tiene la aplicación del submuestreo balanceado o desbalanceado dentro de las unidades o parcelas experimentales en varias disciplinas del conocimiento, como en el caso de las ciencias agropecuarias y forestales, pero se ha encontrado que existe poca información publicada cuando se aplica un diseño experimental, particularmente cuando se analizan datos provenientes de ensayos factoriales. En esta investigación se generan algunos modelos estadísticos para analizar datos en un arreglo de parcelas divididas en diseño experimental en cuadro latino, sin y con submuestreo balanceado; también se construyen formulas generales para calcular grados de libertad y suma de cuadrados, en esta última modalidad con base en la metodología de mínimos cuadrados y formulas cuadráticas o matriciales; se presenta un ejemplo para obtener un análisis de varianza y una comparación de medias de tratamientos con la prueba de la diferencia mínima significativa honesta, mejor conocida como prueba de Tukey; adicionalmente, se analizan datos artificiales para mostrar cómo debe aplicarse un paquete estadístico, como InfoStat e InfoGen. Por la relevancia que presenta el Sistema para Análisis Estadístico (Statistical Analysis System, SAS) en el contexto mundial, sería muy deseable realizar la validación de los procedimientos que se presentan en este estudio aplicando dicho paquete estadístico. Debido a que tampoco se ha reportado información con relación a la serie de experimentos para este tipo de arreglo de unidades experimentales en el diseño experimental de referencia, en este estudio sólo fue generado el modelo estadístico y los procedimientos algebraicos que permitirán obtener un ANAVA y una prueba de Tukey, en ambos softwares referenciados previamente, para el caso general de sin submuestreo balanceado. Es decir, quedará pendiente para otra investigación la parte metodológica que tiene que ver con el caso de submuestreo balanceado, así como su

aplicación agropecuaria y forestal. También es recomendable generar los procedimientos que permitan obtener ambos análisis estadísticos para un arreglo de parcelas subdivididas en Cuadro Latino, sin y con submuestreo.

Palabras clave: Submuestreo, Ensayos factoriales, Suma de Cuadrados, Paquete Estadístico.

I. INTRODUCCIÓN GENERAL

El propósito de la experimentación agrícola, pecuaria y forestal es recolectar información confiable que facilite la comparación entre distintos tratamientos y la validación o negación de hipótesis de trabajo. Fernández *et al* (2010) mencionan que el proceso experimental incluye diferentes fases hasta su finalización entre las que se encuentran: la identificación del problema a solucionar, la definición de las metas, la elección adecuada de los tratamientos a implementar, del material vegetal a utilizar en el experimento, el diseño experimental y el arreglo del mismo, la recolección adecuada de datos, su análisis y la correcta interpretación y exposición de los resultados. Por lo tanto, no basta con recolectar y exponer datos, sino que se deben adquirir correctamente y darle un significado a los mismos. El elemento estadístico es solo una sección, a pesar de su relevancia, de la experimentación; es el instrumento eficaz para el diseño adecuado del experimento y el análisis de los datos obtenidos.

Para determinar si un tratamiento tuvo algún efecto, las hipótesis estadísticas ayudan a determinar tal propósito. Regularmente se miden y cuantifican dos o más especies dentro de cada unidad experimental cuyo fin es mejorar e incrementar la precisión con las que se examinan tales hipótesis (Little y Jackson, 2001).

Zamudio y Alvarado (1996) mencionaron la importancia que tiene el valor de la variable-respuesta que permite obtener el efecto de los tratamientos, la cual se cuantifica en las muestras que pueden ser las unidades experimentales completas o una porción aleatoria de las mismas; en ambos casos que se mencionan, se considera una sola observación por cada unidad experimental. Sin embargo, por razones prácticas o para aumentar la precisión del experimento se pueden tomar dos o más observaciones por unidad experimental; a este hecho se le conoce como submuestreo y para hacer el análisis de los datos experimentales se deberá considerar una

componente de submuestreo en el modelo lineal, un algoritmo para llevar a cabo submuestras en diseños experimentales utilizando el álgebra matricial (Zamudio y Alvarado, 1996). De esta manera los usuarios pueden considerar cada una de las observaciones disponibles dentro de cada parcela o unidad experimental (Gómez y Gómez, 1984; Martínez, 1988; Zamudio y Alvarado, 1996).

Cuando hay que estudiar de forma simultánea los efectos que tienen varios factores sobre una variable respuesta, se diseñan experimentos factoriales en los que todos los niveles de un factor se combinan con todos los niveles de cualquier otro factor para formar una estructura de tratamientos. Así, en un experimento de tres factores, como A, B y C, con a, b y c niveles cada uno, respectivamente, el número total de tratamientos será abc , en r repeticiones.

Los arreglos de unidades experimentales en Parcelas Divididas (PD) o Subdivididas (PSD) han sido utilizados ampliamente para estudiar efectos o varianzas en ensayos bi o trifactoriales (Balzarini *et al.*, 2008; Di Rienzo *et al.*, 2008; Balzarini y Di Rienzo, 2016; Tirado y Tirado, 2017), o en series de experimentos cuando éstos se extienden a diversos ambientes, conformados por años, localidades o combinaciones de ambos (González *et al.*, 2019; Pérez *et al.*, 2022).

El diseño en Cuadro Latino (DCL), en PD o PSD, no ha sido muy utilizado, pero podría ser de gran utilidad para evaluar las diferencias que originan años, localidades o ambos con distanciamiento entre plantas, fórmulas de fertilización, abonos orgánicos, insecticidas, fungicidas, herbicidas, métodos de labranza, láminas de riego, y fechas de corte en forrajes, entre otros, cuando éstos se asignen a parcela principal, y en subparcela podrían considerarse especies vegetales o variedades contrastantes de alguna de éstas (González *et al.*, 2019; Pérez *et al.*, 2022).

De acuerdo a lo anterior se tuvo como:

Hipótesis

El uso de submuestreo balanceado dentro de unidades o parcelas experimentales bajo un diseño experimental es igualmente viable con relación al empleo de medias o promedios aritméticos, pero será más confiable cuando el número de plantas o unidades de observación dentro de cada una de estas sea mayor debido a la correlación positiva que existe entre tamaño de muestra y precisión.

Y como:

Objetivo General

Construir los modelos estadísticos con y sin submuestreo pertinentes a la presente investigación.

Objetivos Específicos

- Aplicar submuestreo balanceado dentro de cada unidad experimental considerando datos artificiales, para validar las fórmulas vinculadas a dichos modelos por medio de la generación de un Análisis de varianza y una comparación de medias de tratamientos.
- Diseñar un procedimiento general para construir bases de datos y analizar éstos aplicando InfoStat o InfoGen.

II. REVISIÓN DE LITERATURA

2.1 Definición de conceptos

2.1.1 Submuestreo

El submuestreo es una técnica que es utilizada en los experimentos agrícolas, pecuarios y forestales que se establecen en campo, en invernadero, en la agroindustria o en un bosque, entre otros, para recolectar información de los tratamientos que están siendo evaluados en las unidades experimentales a partir de dos o más observaciones con la finalidad de analizarlos estadísticamente (Zamudio y Alvarado, 1996).

Para Fernández *et al* (2010) la meta de la experimentación es conseguir información confiable que permita realizar comparaciones entre distintos tratamientos y respaldar o descartar hipótesis de trabajo. Esto se conseguirá con la exactitud con la que sean recolectados. El procedimiento experimental abarca varias fases, desde la identificación del problema a resolver, la definición de los objetivos, la elección adecuada de los tratamientos a implementar, del material vegetal a utilizar, el diseño experimental, la recolección adecuada de datos, su análisis y la correcta interpretación y exposición de los resultados. No es suficiente con recolectar y exponer datos, sino que se deben adquirir correctamente y encontrarle un significado. El componente estadístico constituye solo un segmento, a pesar de ser relevante.

El uso de “S” datos que se registrarán dentro de cada Unidad experimental (UE) contribuye a pruebas de hipótesis estadísticas más precisas con relación a la estructura de los tratamientos que están siendo evaluados cuando más plantas o unidades de observación son medidas o registradas entro de cada una de éstas. Para Gavilánez (2024), la precisión en los experimentos es un elemento crucial para identificar diferencias entre los diferentes tratamientos evaluados

en un ensayo. Su mayor ocurrencia se basa fundamentalmente en un diseño adecuado, una correcta gestión experimental y un procedimiento pertinente para la medición de las variables respuesta. En este contexto, un experimento exacto contribuye a cometer menos errores como negar una opción que en un contexto específico tiene efectos reales (error tipo II), no obstante, debido a la ausencia de control sobre la varianza interna del ensayo, estos son evidentes a nivel experimental. Esta falta de exactitud genera así un falso negativo que, potencialmente, desechará alguna posibilidad de incrementar la productividad de alguna cosecha o, incluso, de algún cultivo.

Little y Hills (2001) definen a una submuestra como un conjunto de mediciones que constituye parte de una población, a partir de la cual obtenemos información y hacemos inferencias acerca de la misma. Por tal motivo, es importante que la submuestra sea representativa de la población que se está considerando. Para que sea representativa se utilizan técnicas aleatorias de muestreo, donde cualquier medición individual tiene tantas probabilidades de ser incluida como cualquier otra. Para mejorar dicha precisión es indispensable el uso de submuestreo para todas las variables cuantitativas que han sido registradas (usar cada una de las “S” observaciones disponibles dentro de cada parcela o UE) como lo proponen Gonzalez *et al* (2023).

2.1.2 Diseño experimental y factorial

El diseño de un experimento implica la determinación o modificación intencionada de alternativas a evaluar en un proceso productivo como el agrícola (o cualquier otro), bajo normas estrictas que garanticen que la variación que generen estas alternativas de prueba no sea resultado de factores aleatorios (Gavilánez, 2024).

Esta metodología del método científico tiene como objetivo transitar del enfoque deductivo al inductivo, proporcionando sugerencias generales para los agricultores que les faciliten incrementar su productividad y rentabilidad. Para alcanzar este objetivo, el diseño de un experimento debe ser eficiente en términos de controlar el error experimental y efectivo para verificar las hipótesis propuestas (Pantoja, 1989); esto implica que se pueda obtener la mayor información posible de un ensayo con la mínima inversión económica, considerando el costo de una investigación y las restricciones económicas que siempre existen.

Se deben considerar para su óptimo desarrollo tres principios esenciales para satisfacer las exigencias del método científico presentes en los ensayos en general, además de tener en cuenta el protocolo que cada uno requiere. El primero de ellos es la aleatorización en la distribución de los tratamientos a las unidades experimentales (o viceversa), lo que permite la autonomía de los residuales en los modelos estadísticos. El segundo es la repetición de los tratamientos, sin el cual resulta inviable evaluar de manera eficaz el error experimental. Por último, el tercer principio es introducir el principio de bloqueo (formación de bloques) en el experimento, también conocido como control local. Se refiere a distribuciones que se aplican en ambientes no controlados, tales como las de bloques completos e incompletos al azar o cuadrados latinos que se utilizan en la experimentación agrícola o de campo; y por su exigencia de homogeneidad en las condiciones de suelo, no rige para las distribuciones completamente al azar, situación ante la cual una distribución como la de bloques al azar permitirá tener un mejor control de esta fuente de variación secundaria, reduciendo la influencia de dicho efecto al formar pequeños bloques homogéneos, de tal forma que esta variabilidad se vea reflejada entre dichos bloques y no dentro de los tratamientos (Gavilánez, 2024; Gómez y Gómez, 1984; Little y Jackson, 2001; Martínez, 1994).

Cuando se necesita examinar al mismo tiempo los impactos de varios factores en una variable respuesta, se emplean experimentos factoriales. En estos, todos los niveles de un factor se combinan con todos los niveles de cualquier otro factor para construir una estructura de tratamientos. Por lo tanto, en un experimento factorial con a, b y c niveles, respectivamente, la cantidad total de tratamientos es abc (Fernández *et al*, 2010).

Una de las restricciones principales de los experimentos factoriales es que al aumentar los factores aumenta también el número de interacciones y su complejidad, lo que hace más difícil la interpretación de los resultados; con ocho factores de la productividad y dos niveles dentro de cada uno de éstos habrá un total de 247 interacciones (Pérez *et al.*, 2019).

Fernández *et al* (2010) y Gómez y Gómez (1984) comentaron que a través de los experimentos múltiples se hacen inferencias hacia uno o más factor, así como para sus interacciones posibles utilizando alguno de los siguientes diseños experimentales: completamente aleatorio, bloques completos al azar, cuadrado latino, o látice. No obstante, cuando se incrementa significativamente el número de factores y niveles de estos, ciertos diseños, como el cuadrado latino, requiere de un número de repeticiones que lo hacen inviable, pero también en ocasiones, es complicado conseguir repeticiones uniformes en un diseño de bloques completos al azar, lo que provoca que el error experimental se incremente. Por ello, cuando el número de tratamientos es elevado, se utilizan arreglos para estos diseños, como el de bloques incompletos que sacrifica las interacciones de alto orden, o el de parcelas divididas o subdivididas que sacrifican un factor o un grupo de factores y sus interacciones, o bien, otros diseños en los que la pérdida de precisión se distribuye uniformemente entre los factores y sus interacciones, como látices balanceados, bloques incompletos balanceados y látices cuadrados (Cochran y Cox, 1978; Gomez y Gomez, 1984; Martínez, 1988).

2.1.3 Diseño Cuadro o cuadrado latino

El nombre de cuadro latino se deriva del hecho de que los tratamientos se representan, usualmente, por letras del alfabeto latino. La característica fundamental del diseño, consiste en que un tratamiento cualquiera aparece representado una sola vez en la misma hilera y en la misma columna. Las unidades experimentales se agrupan en bloques completos, de acuerdo con los dos criterios de clasificación del material experimental, uno siendo perpendicular al otro (Martínez, 1994).

En experimentos de campo permite que se elimine la variación que proviene de las diferencias de fertilidad en el suelo en dos direcciones cruzadas, o bien la variación de la fertilidad del suelo en una dirección y el gradiente de humedad generado al dar los riegos en otra dirección perpendicular. Puede ser útil también en experimentos de invernadero en donde la iluminación y la temperatura pueden ser dos fuentes de variación evidentes que deberían extraerse del error experimental o residual del modelo (Fernández *et al*, 2010; Gomez y Gomez, 1984; Martínez, 1988).

Para Gómez y Gómez (1984) este es un diseño el cual tiene la capacidad para manejar simultáneamente dos fuentes conocidas de variación dentro y entre unidades experimentales. Trata las fuentes como dos criterios de bloqueo independientes, en lugar de solo uno como en el diseño de bloques completos al azar. La presencia de bloqueo de filas y bloqueo de columnas, si bien es útil para atender dos fuentes independientes de variación, también se convierte en una restricción importante en el uso del diseño. Esto es así porque cada tratamiento aparece una sola vez en cada bloque de filas y en cada bloque de columnas, y por lo tanto, debe satisfacerse el requisito sólo si el número de replicaciones es igual al número de tratamientos. Por ejemplo, en un cuadro latino 10 x 10, los 10 tratamientos serán evaluados en 10 repeticiones. Otra

consideración importante es que cuando el número de tratamientos es grande, el diseño se vuelve impráctico debido al gran número de repeticiones requeridas, resultando excesivo en costos y poco práctico. Por otro lado, cuando el número de tratamientos es pequeño, los grados de libertad asociado con el error experimental se vuelve demasiado pequeño para que el error sea confiable; en la práctica frecuentemente se diseñan experimentos en cuadrados latinos eligiendo entre cuatro y ocho tratamientos. Otra desventaja del diseño de cuadro latino es que el número de filas, columnas y tratamientos debe ser el mismo, y el análisis también se complica si se pierden unidades experimentales en el transcurso del experimento (Fernández *et al*, 2010; Gomez y Gomez, 1984; Martínez, 1988).

2.1.4 Parcelas divididas o “split-plot Design”

En éste las unidades o parcelas principales, a las que se les aplican los niveles de uno o más factores, se dividen en subunidades o subparcelas, a las cuales se les asignan los niveles de otro o otros factores. De esta manera, cada unidad principal completa se convierte en un bloque para los niveles asignados a las subunidades (Fernández *et al*, 2010).

Para el Dr. Ángel Martínez Garza (1994) este arreglo en la experimentación es apropiado para aquellas situaciones en que un factor es difícil de ensayar en parcelas pequeñas, como ocurre con los métodos de cultivo que se aplican con maquinaria pesada, en los experimentos de riego, etc., y no debería usarse, en aquellos casos donde el tamaño de la parcela sea independiente del tratamiento aplicado.

Para Little y Jackson (2001) los diseños de parcelas divididas se emplean frecuentemente en experimentos factoriales, en los que la naturaleza del material o las operaciones contempladas dificultan el manejo de todas las combinaciones de factores en una misma forma. En éstos se

suele sacrificarse la precisión en la estimación de los efectos promedio de los niveles asignados a las parcelas principales, aunque frecuentemente aumenta la precisión para comparar los efectos promedio de niveles asignados a subparcelas; y cuando existen interacciones, para comparar los efectos de tratamientos de subparcelas en un tratamiento de una parcela principal dada.

Por su naturaleza factorial donde se ven involucrados dos factores en un experimento, Gómez y Gómez (1984) mencionan que este tipo de arreglo es ideal cuando existen más tratamientos de los que pueden acomodarse en un diseño de bloques completos al azar.

Fernández *et al* (2010) mencionaron que este arreglo de unidades experimentales confina la exactitud en la estimación de los impactos del factor principal, aunque generalmente acrecienta la exactitud para contrastar los efectos del factor subordinado y los de su interacción con el factor principal. Esto ocurre debido a que, al tener dos tamaños de parcelas, se presentan dos errores experimentales. Usualmente, el error de las subparcelas es inferior al de las parcelas principales.

Para Gómez y Gómez (1984) la asignación de un factor particular a la parcela principal o a la parcela secundaria es extremadamente importante, debido a que en parcelas divididas, el tamaño de la parcela y la precisión de la medición de los efectos no son los mismos para ambos factores. Para hacer tal elección, ellos hacen las siguientes recomendaciones:

- a) **Grado de Precisión.** Para obtener un mayor grado de precisión para el factor B que para el factor A, asigne el factor B a la parcela secundaria.
- b) **Tamaño relativo de los efectos principales.** Si se espera que el efecto principal de un factor (factor B) sea mucho mayor y más fácil de detectar que el del otro factor (factor

A), el factor B puede asignarse a la parcela principal. Esto aumenta las posibilidades de detectar la diferencia entre los niveles del factor A, que causen un efecto menor.

- c) **Prácticas de Gestión.** Las prácticas culturales requeridas por un factor pueden dictar el uso de grandes parcelas. Por conveniencia práctica, dicho factor puede asignarse a la parcela principal.

La cantidad de factores ponderados que participan en estos diseños puede exceder a más de dos, generando así otras variantes del mismo diseño. Por lo tanto, cuando se tienen tres factores se denomina parcelas subdivididas o “split-split-plot Design”. En estas circunstancias, contaríamos con tres tamaños de parcela (principal, subparcela y sub-subparcela) y, en consecuencia, tres errores experimentales. También se denominan parcelas grandes, medianas y chicas, respectivamente. Sin embargo, el incremento en la cantidad de factores jerarquizados dificulta significativamente las interacciones, los cálculos y la interpretación de los resultados, lo que los hace menos empleados. En el diseño de parcelas divididas, tanto el procedimiento de aleatorización como el de análisis de varianza se lleva a cabo en dos etapas: una en el nivel de la parcela principal y otra en el nivel de la subparcela (Gomez y Gomez, 1984; Martínez, 1988; Tirado y Tirado, 2017).

2.1.5 Serie de experimentos

Para Gómez y Gómez (1984) la productividad de los cultivos se basa en el genotipo, el ambiente en el que éstos se desarrollan y en la interrelación entre el genotipo y el entorno (ambiente). En la práctica científica se tiene la capacidad de regular el genotipo y ciertos elementos ambientales, tales como la cantidad de fertilizante, la cantidad de plantas y el control de plagas. Sin embargo, otros elementos del entorno, como la radiación solar, las precipitaciones pluviales y ciertas características edáficas, suelen ser fijos y complicados de alterar para un lugar y una temporada

de siembra específica. Por lo tanto, en un experimento singular en un solo lugar, se pueden modificar y evaluar únicamente los factores controlables, pero no los elementos ambientales que están fuera de su dominio.

El efecto de los factores ambientales incontrolables sobre el rendimiento de los cultivos es tan importante, si no más, que el de los factores controlables; y la evaluación y cuantificación de sus efectos son igualmente esenciales. Debido a que se espera que los factores incontrolables cambien con la estación y el sitio, y debido a que estos cambios son mensurables, se pueden evaluar sus efectos sobre el desempeño del tratamiento. En la experimentación agrícola, la forma más comúnmente utilizada para evaluar los efectos de los factores ambientales que no se pueden controlar sobre la variable respuesta del cultivo es repetir el experimento en varios sitios, o durante varias temporadas de cultivo, o ambas (Gómez y Gómez, 1984).

Martínez (1994) nos hace mención que fue Gosset (1936) quien por primera vez llamó la atención a los investigadores sobre este tema, en el cual se observa que el efecto de una variedad dada de algún cultivo, puede diferir con el ambiente, y que las conclusiones generales aplicables a una gran región, tienen que basarse en un esfuerzo cooperativo de varias estimaciones experimentales.

Fernández *et al* (2010) mencionaron que un mismo experimento se puede llevar a cabo varias veces en distintos periodos, lugares o años para analizar la adaptación de una tecnología reciente a distintas condiciones ambientales, distintas localidades, o su impacto en la productividad a través del tiempo, o bien, para verificar la repetitividad de los resultados logrados. En cualquier caso, se busca la combinación de información de la variedad de estos experimentos para su análisis conjunto.

Gavilánez (2024) mencionó que la serie de experimentos son estudios longitudinales (prolongados), los cuales buscan confirmar el efecto de algún tratamiento sobresaliente en las primeras pruebas, midiendo su capacidad en ensayos subsecuentes para, de esta forma, valorar su respuesta bajo distintas modificaciones de un mismo clima; a pesar de que también habría condiciones físicoquímicas del suelo que podrían variar de un año a otro.

Existe una variedad de experimentos que pueden valorarse a través del tiempo. Dentro de estos casos están los ensayos de variedades, fórmulas con fertilizantes, plaguicidas, métodos de preparación de suelos, entre otros. Gómez y Gómez (1984) refieren una clasificación de la serie de experimentos, dependiendo el objetivo que se persiga:

- a) **Experimentos de evaluación preliminar.** Diseñados para identificar, desde un punto de vista gran cantidad de nuevas tecnologías, aquellas que brindan un desempeño superior en el área donde se desarrollan.
- b) **Experimentos de adaptación tecnológica.** Diseñados para determinar el rango de adaptabilidad geográfica de las pocas tecnologías superiores identificadas en uno o más experimentos de evaluación preliminar.
- c) **Experimentos a largo plazo.** Diseñados para caracterizar una nueva tecnología con respecto a su efecto a largo plazo sobre la productividad.
- d) **Experimentos de predicción de respuesta.** Diseñados para identificar una relación funcional entre la respuesta del cultivo y algunos factores ambientales para predecir la productividad en una amplia gama de entornos.

Según el número de ensayos llevados a cabo durante varios años, estas experiencias se utilizarán para validar la eficacia de una opción de mejora en cualquier temporada y/o determinar en qué periodos es más aconsejable su implementación. El efecto del clima en los experimentos que se

replican durante varias temporadas podría hacerse presente en el efecto de los bloques (repeticiones) cuando se trata de distribuciones como en los diseños experimentales DBCA o DCL y en la respuesta de los mismos tratamientos. Además, es comprensible que el factor temporada (tiempo) no puede ser aleatorizado como si se tratara de un ensayo factorial sin restricciones, no obstante, la interacción de interés para verificar si la respuesta de los tratamientos es generalizada entre temporadas es la de estas dos fuentes de variación (Gavilánez, 2024).

Martínez (1994) recomienda que cuando se tiene la oportunidad de planear una serie de experimentos, debe buscarse la solución más sencilla de diseño. Cuando se maneja un conjunto de tratamientos que puedan alojarse convenientemente en Bloques Completos al Azar, cada experimento debe comprender todos los tratamientos objeto de estudio. El análisis combinado de los experimentos será más sencillo. Los experimentos planeados, en una buena experimentación, deberán ser independientes; deberán realizarse aleatorizaciones independientes para cada experimento. Con ello se evitará introducir sesgos en las estimaciones e invalidar los niveles de significancia y el poder de la prueba del análisis de varianza. Cualquier desbalance en el diseño de una serie de experimentos implica complicar el método de análisis.

2.2 Importancia de la investigación

En la actualidad existe poca información publicada sobre este tema. Esta situación se ha observado específicamente en el programa estadístico Sistema para Análisis Estadístico (Statistical Analysis System, SAS), pero para InfoStat e InfoGen aun en sus manuales del usuario no existen reportados los procedimientos que permitan estimar el error muestral, componente importante en los modelos estadísticos relacionados con los diseños experimentales DCA, DBCA y DCL, particularmente cuando éstos se extienden a ensayos factoriales con

diversos arreglos de unidades experimentales como las series de experimentos en parcelas divididas, parcelas subdivididas y franjas o bloques divididos, entre otros (Zamudio y Alvarado, 1996; Rodríguez *et al.*, 2023; Rodríguez *et al.*, 2025).

La estimación de efectos y varianzas relacionadas con la estructura de tratamientos en un experimento es prioritaria para probar hipótesis estadísticas más precisas, especialmente cuando se recurre al muestreo dentro de las unidades o parcelas experimentales, de ahí la necesidad de desarrollar los procedimientos correctos que permitan contribuir a dicho propósito, de manera fácil, accesible y segura (Gomez y Gomez, 1984; Martínez, 1988).

Ojeda y Gutiérrez (2005) comentaron que hay poca información sobre la reformulación de los modelos lineales jerárquicos actualmente en el contexto de series de experimentos, a pesar de que muchos resultados del área del meta análisis pueden relacionarse directamente, y de que algunas formulaciones particulares de estos modelos existen desde la década de los años cincuenta del siglo pasado. Por lo tanto, se puede afirmar que esta metodología necesita un análisis más detallado y profundo sobre su viabilidad para el estudio de datos de experimentos individuales o de series de ensayos (Zamudio y Alvarado, 1996; Gomez y Gomez, 1984; Martínez, 1988; Rodríguez *et al.*, 2023).

2.3 Modelaje

El diseño fundamentado en las unidades experimentales se distingue por no ser una disposición geométrica concreta de los tratamientos realizados en el campo. Es un diseño conceptual que permite visualizar cualquier diseño experimental; esto implica que cualquier estructura geométrica de los tratamientos aplicados a las unidades experimentales, puede ser presentada como un diseño de las mismas (Zamudio y Alvarado, 1996). Se le nombra modelo lineal o

matemático, y está asociado con el Análisis de Varianza (ANOVA), por medio del cual podemos entender la variación que está presente dentro del experimento realizado (Fernández *et al*, 2010; Martínez 1994), y se conforma por los siguientes elementos:

$$y_{ij} = \mu + U_i + \varepsilon_{ij}$$

Donde:

y_{ij} representa el valor observado de la característica en estudio sobre la unidad experimental de un determinado experimento; μ es el efecto general de la media; U_i es el efecto aleatorio general controlado de la i -ésima unidad experimental; ε_{ij} es el efecto aleatorio no controlado de dicha observación dentro de la unidad experimental a la que pertenece, que es el error experimental o residual de ese modelo.

Este ejemplo de modelo lineal puede extrapolarse para los diferentes diseños experimentales. Pérez *et al* (2024) propusieron el siguiente modelo para un experimento en cuadro latino en arreglo de parcelas divididas:

$$Y_{ijklm} = \mu + E_m + H_i + C_j + A_k + (EA)_{mk} + (EHA)_{i(mk)} + B_l + (EB)_{ml} + (AB)_{kl} + (EAB)_{mkl} + \varepsilon_{ijklm}$$

Dónde: Y es la variable que será analizada; μ es la media aritmética general; E_m , H_i , C_j , A_k , B_l , son los efectos principales causados por los factores identificados como Ambientes, Hileras, Columnas, A, B, respectivamente; $(EHA)_{i(mk)}$ es el Error a; $(AB)_{kl}$, $(EA)_{mk}$, $(EB)_{ml}$, $(EAB)_{mkl}$ son interacciones de primer y segundo orden originadas por la combinación de dos o tres de los factores previamente descritos; ε_{ijklm} es el Error b o residual del modelo.

Para efectos del presente trabajo de investigación se desarrollaron dos modelos, para un solo experimento en cuadro latino con arreglo en parcelas divididas, con y sin submuestreo. Ambos modelos se construyeron aplicando la guía publicada por Sahagún (1998), y sus componentes son:

Sin submuestreo:

$$Y_{ijkl} = \mu + H_i + C_j + A_k + (HA)_{ik} + B_l + (AB)_{kl} + \varepsilon_{ijkl}$$

Con submuestreo:

$$Y_{ijklm} = \mu + H_i + C_j + A_k + (HA)_{ik} + B_l + (AB)_{kl} + \delta_{m(ikl)} + \varepsilon_{ijklm}$$

Dónde: Y es la variable respuesta; μ es la media general; H_i , C_j , A_k , B_l , son los efectos causados por hileras, columnas, Factor A y Factor B, respectivamente; $(HA)_{ik}$ es el error a de la parcela grande; $(AB)_{kl}$ es la interacción originada por los niveles de ambos factores; $\delta_{m(ikl)}$ es el error muestral; ε_{ijkl} y ε_{ijklm} son el error experimental, sin y con submuestreo, correspondientes a las parcelas chicas o subparcelas; éstos últimos también pueden considerarse como residuales en ambos modelos estadísticos.

2.4 El análisis de varianza

El Análisis de la Varianza (ANOVA) es una técnica estadística altamente eficaz para analizar el efecto de uno o más factores sobre la media o promedio aritmético de una variable respuesta. La premisa fundamental del ANOVA radica en descomponer la variabilidad total observada en unos datos en una serie de términos, vinculados a los impactos de cada factor estudiado y a sus posibles interacciones, más una parte residual con la que después se realizarán las pruebas de F (Gomez y Gomez, 1984; Martínez, 1988; Sahagún, 1993).

El análisis de varianza frecuentemente utiliza las medias de los agrupamientos o clasificaciones que se generan de las unidades experimentales, en los diferentes diseños experimentales. Los arreglos de las unidades experimentales se clasifican según la estructura de los tratamientos; en algunos casos, la clasificación preferente es en bloques, filas, parcelas principales y otras modalidades. El análisis de varianza utiliza las medias de dichos agrupamientos, denominadas fuentes de variación, para estimar efectos o varianzas o cuadrados medios (Little y Jackson, 2001). Para Gómez y Gómez (1984) en la situación más sencilla cuando existen dos fuentes de variación entre las n observaciones obtenidas de un experimento, uno corresponde a la variación del tratamiento, el otro es el error experimental. El tamaño relativo de los dos se utiliza para indicar si la diferencia observada entre tratamientos es real o se debe al azar. Se dice que la diferencia de tratamientos es real si la variación del tratamiento es suficientemente mayor que la correspondiente al error experimental o residual del modelo.

Para Cochran y Cox (1978) el ANAVA es una técnica que ayudará a acelerar la obtención de suma de cuadrados de las observaciones involucradas en las unidades experimentales que fueron tratadas. El modelo propuesto inicialmente está compuesto por la suma de cuatro componentes: la media general, efecto del tratamiento, efecto del ambiente y efecto del residual del modelo; en éste se dividen las sumas de cuadrados en cuatro componentes, una atribuible a la media

general, una a las diferencias entre la estimación de los efectos de los tratamientos y una a los efectos ambientales que el experimento es capaz de medir, y por último, una que es el residual o suma de cuadrados de los errores. En la mayor parte de los casos, se calcula la suma de cuadrados original y las tres primeras componentes, obteniendo la suma de cuadrados del error, por sustracción. El análisis de varianza ofrece más que un método abreviado para obtener la suma de cuadrados del error. La suma de cuadrados debida a los tratamientos es la cantidad necesaria para la prueba F de la hipótesis de que no existen diferencias entre los efectos de los tratamientos. Con una pequeña extensión, el análisis también da la suma de cuadrados requerida para probar la igualdad de los efectos de un subgrupo de los tratamientos. La componente debida a los efectos ambientales permite estimar en cuanto aumenta la exactitud del experimento, eliminando estos efectos de las estimaciones de las medias de los tratamientos.

Zamudio y Alvarado (1996) mencionaron que cuando se habla de submuestreo en diseños experimentales, se hace referencia a que en cada unidad experimental a las que se asignan los tratamientos, se toman dos o más observaciones llamadas submuestras; y la suma de cuadrados total (SCT), independientemente del diseño experimental, queda conformada por la suma de cuadrados de las observaciones dentro de las unidades experimentales (SCEM) adicionada a la suma de cuadrados de las unidades experimentales (SC(UE)). Esta partición de la SCT, constituye el punto central para el análisis de cualquier diseño experimental con submuestreo, y se puede representar simbólicamente como:

$$SCT = SC(UE) + SCEM$$

2.5 Comparación de medias de tratamientos

Cuando en el análisis de varianza resulta estadísticamente significativa alguna fuente de variación de interés, por ejemplo, la relacionada con la estructura de los tratamientos, el siguiente paso consiste en indagar entre cuales de ellos hay diferencias reales en la variable respuesta. El propósito de la experimentación consiste en responder preguntas que se plantean al principio a manera de hipótesis, donde por medio de la obtención de las medias correspondientes podría inferirse de manera estadística el rechazarla o aceptarla.

Little y Jackson (2001) hacen mención de algunos principios a considerar para decidir sobre los tratamientos y qué estos pueden ayudarnos a seleccionar un plan para separar las medias. Entre estos se encuentran:

- **El arreglo factorial.** Cuando se han de estudiar dos o más factores simultáneamente, el arreglo factorial hace posible el estudio de las interacciones entre factores. Cuando las interacciones no están presentes, los resultados son más ampliamente aplicables, puesto que los tratamientos principales se prueban para un rango de condiciones más amplio.
- **La selección de niveles de tratamiento para obtener información sobre tendencias o respuestas óptimas.** Cuando los tratamientos consisten en dosis, como fertilizantes a base de nitrógeno, fósforo y potasio, la pregunta fundamental no es si alguna dosis resulta significativamente mejor que otras, si no si existe o no una respuesta al factor estudiado y cómo se caracteriza mejor esa respuesta. Una atención cuidadosa en la selección de los niveles de dosificación, es importante para caracterizar una función dosis-respuesta.
- **La utilización de un tratamiento estándar.** Puede resultar deseable comparar cada uno de los diversos tratamientos con un tratamiento estándar o control a fin de decidir si

cualesquiera de los nuevos tratamientos son mejores que el que se está utilizando actualmente.

Fernández *et al* (2010) mencionaron que los procedimientos para contrastar grupos de medias después del análisis de varianza se les designa comparación de contrastes y éstos se pueden clasificar en dos categorías principales: contrastes *a priori* y *a posteriori*. Como su nombre lo indica, los primeros son los que se definen antes de iniciar de llevar a cabo el experimento; los segundos son todas las posibles comparaciones de medias en parejas diferentes o en grupos de ellas cuando el experimento ya fue concluido. No obstante lo anterior, para evitar sesgo estadístico en su aplicación, deben diseñarse contrastes mutuamente ortogonales, debido a que en éstos existen $t-1$ comparaciones que son únicos y mutuamente excluyentes (Gomez y Gomez, 1984; Martínez, 1988; Pérez *et al.*, 2019).

Los cuantiosos métodos utilizados para realizar comparaciones entre las medias de los tratamientos varían principalmente en función del nivel de protección frente a los dos tipos de error asociados a cualquier contraste de hipótesis estadística: tipo 1 (falsos positivos, acierto H_0) y tipo 2 (falsos negativos, acierto H_1). Los dos tipos de error están inversamente relacionados, luego los métodos que den más protección frente al error tipo I ofrecerán menor protección frente al error tipo II y viceversa. Para Little y Jackson (2001) el problema de decidir que medias del tratamiento son significativamente diferentes reciben el nombre de separación de medias. Existen tres enfoques generales para su separación: la aplicación de las diferencias significativas mínimas; la utilización de las pruebas de rango múltiple, y las pruebas de F *a priori*.

Los métodos más usados, según Fernández *et al* (2010), para comparar medias según el tipo de comparación pueden ser: *a priori* o planificadas y *a posteriori*.

2.5.1 Comparaciones *a priori*

Se utilizan en experimentos que tienen una estructura de tratamientos establecida de antemano, es decir, las comparaciones a realizar son hipótesis especificadas antes de realizar el experimento. Esta es la situación deseable siempre porque reduce el número de comparaciones, dejando sólo las esenciales y eliminando las irrelevantes, reduciendo así el riesgo de cometer errores tipo I y II. El método más utilizado en factores cualitativos, como por ejemplo la elección de un grupo de variedades de maíz (*Zea mays* L.), es el de los contrastes ortogonales o mutuamente ortogonales, aunque también son empleados los polinomios ortogonales cuando se estudia el efecto que causa un factor cuantitativo sobre una variable respuesta, por ejemplo el incremento en la altura de las plantas de papa (*Solanum tuberosum* L.) debido a la aplicación de nitrógeno en el intervalo de 0 a 400 kg por ha (Gomez y Gomez, 1984; Martínez, 1984; Martínez, 1988; Pérez *et al.*, 2019).

Contrastes ortogonales. La estrategia de este método consiste en dividir la suma de cuadrados de los tratamientos entre las diversas hipótesis o contrastes relacionadas con $t - 1$ comparaciones, donde t es el número de tratamientos que está siendo evaluado. La suma de cuadrados asociada a cada contraste se usa como numerador y el cuadrado medio del error como denominador para calcular razones F y comprobar su significación (valor P). Cada contraste, por lo general, consume un grado de libertad, por lo que tendremos $t-1$ contrastes diferentes que son los grados de libertad correspondientes a los tratamientos. Si los contrastes son mutuamente ortogonales, la adición de las sumas de cuadrados de los $t-1$ contrastes de este tipo será igual a la suma de cuadrados de tratamientos (Gomez y Gomez, 1984; Martínez, 1984; Martínez, 1988; Pérez *et al.*, 2019).

Polinomios ortogonales. Cuando se analiza un factor cuantitativo (temperatura, tiempo, densidad de inóculo, dosis de fertilizantes, distanciamiento entre plantas, entre otros), el concepto de contrastes ortogonales se puede extender a los diferentes niveles de este factor para determinar si existe alguna tendencia (lineal, cuadrática, cúbica) en la respuesta. En este caso, los procedimientos de comparación múltiple de medias son claramente inapropiados. El contraste entre polinomios ortogonales tiene su principal aplicación cuando los niveles del factor cuantitativo son pocos (<5). Sin embargo, cuando los niveles de ese factor son mayores, el análisis de regresión es más confiable. Al igual que en los contrastes ortogonales, la suma de cuadrados del factor se divide en diferentes contrastes (lineal, cuadrático, cúbico, etc.), casi siempre con un grado de libertad cada uno. El contraste lineal determina si existen diferencias significativas entre los niveles bajos y los niveles altos del factor cuantitativo. El cuadrático detecta diferencias entre los niveles intermedios y los extremos. El cúbico compara entre niveles alternos (Gomez y Gomez, 1984; Martínez, 1984; Martínez, 1988; Tirado y Tirado, 2017).

La hipótesis nula (H_0) es que no existe ningún efecto significativo que sea de naturaleza lineal, cuadrático, cúbico, etc., debido a los diferentes niveles del factor cuantitativo. Algunos paquetes estadísticos utilizados para estimar contrastes o polinomios ortogonales son SAS, InfoStat, InfoGen, OPSTAT, Agrobase, y WASP, entre otros.

Comparaciones contra un tratamiento control. Cuando se pretende comparar un conjunto de medias, una a una, frente a una única media, la del tratamiento testigo o control, el procedimiento más usado es la prueba de Dunnett; en ésta se establece un valor crítico para la comparación de dos medias, a partir del cual podemos declarar que dichas medias difieren significativamente. Se trata, por tanto, de un método conservador que protege principalmente contra el error tipo I (falsos positivos). No obstante, si se compara con los métodos conservadores de comparaciones

múltiples es más liberal, o potente, que el más liberal de éstos, la de la diferencia mínima significativa honesta, mejor conocida como prueba de Tukey, ya que al haber menos comparaciones disminuye el riesgo de cometer el error tipo II (falsos negativos).

Comparaciones con el mejor. Cuando en lugar de comparar con la media del control se pretende identificar los tratamientos que producen el mejor resultado (media mayor o menor), se utiliza la prueba de Hsu (1996). Al igual que la de Dunnett, la de Hsu es un método conservador pero, al reducir el número de comparaciones diferentes, aumenta su protección contra el error tipo II siendo, por tanto, más potente o liberal que la de Tukey.

2.5.2 Comparaciones *a posteriori*

Ésta se utiliza en experimentos que no tienen una estructura *a priori* en la estructura de los tratamientos y permite realizar todas las comparaciones entre medias aritméticas de parejas diferentes [$t(t - 1) / 2$]. Al ser los procedimientos más empleados tradicionalmente, se han denunciado numerosos abusos derivados de su utilización para comparar medias de un factor cuantitativo o de experimentos con una estructura claramente definida de los tratamientos (Gomez y Gomez, 1984; Little y Hills, 2008).

Los métodos de comparaciones múltiples son numerosos y se pueden agrupar en dos categorías en función de la protección que ofrecen contra los errores de tipo I y II. Los que presentan una mayor potencia contra el error tipo II (falsos negativos) y, por tanto, ofrecen una menor protección contra el error tipo I (falsos positivos), se consideran “liberales”, mientras que los que protegen especialmente contra el error tipo I y presentan una menor potencia contra el error tipo II, serían los “conservadores”. En el primer grupo se incluyen la mínima diferencia significativa (MDS o “least significant difference, LSD”), o mínima diferencia significativa de

Fisher (FLSD), y la LSD de Waller-Duncan-Bayes (BLSLSD). En el segundo grupo se incluyen las pruebas de Duncan, Student-Newman-Keuls (SNK), Tukey, Sidak, Bonferroni y Scheffé (Balzarini *et al.*, 2008; Di Rienzo y Balzarini, 2016).

2.6 Uso de Software

El procesamiento estadístico de datos provenientes de arreglos en Parcelas Divididas (PD) o Subdivididas (PSD) se puede realizar de manera fácil y confiable utilizando InfoStat y InfoGen (Balzarini *et al.*, 2008; Di Rienzo *et al.*, 2008; Balzarini y Di Rienzo, 2016) y SAS, así como el correspondiente a las series de experimentos a través de años, localidades o ambos (González *et al.*, 2019; Pérez *et al.*, 2022; González *et al.*, 2024 b).

Los diseños experimentales que más se han utilizado son Completamente al Azar (DCA) y Bloques Completos al Azar (DBCA) (Martínez, 1988; Gomez y Gomez, 1984; Little y Hills, 2008; Montgomery, 2009), pero Ledolter (2010) analizó y discutió los arreglos en PD para ensayos factoriales y para factoriales fraccionados, con énfasis en los diseños experimentales DCA y DBCA.

La serie de experimentos para un DCL, en PD o PSD podría ser fundamental para ayudar a detectar diferencias importantes entre años, entre localidades o entre ambos cuando se elijan diversos paquetes tecnológicos o diferentes manejos agronómicos, orientados a la generación, validación, aplicación y/o transferencia de tecnología a terrenos de agricultores, con el propósito de incrementar la productividad en los cultivos, y para mejorar la calidad de la materia prima que se deriva de ellos y que será utilizada en la agroindustria mexicana. El mejoramiento genético, la producción de semillas o la hibridación intra e interespecífica también podría documentarse mejor con base en estas metodologías estadísticas (González *et al.*, 2019; Pérez

et al., 2022; González *et al.*, 2024 a, b). Los terrenos de agricultores que son muy heterogéneos también presentan este tipo de variabilidad aleatoria que es indeseable para establecer un experimento adecuado; dicha heterogeneidad ambiental podría controlarse más eficientemente utilizando un PD o un PSD en DCL (González *et al.*, 2019; Pérez *et al.*, 2022; Rodríguez *et al.*, 2025).

También se han encontrado pocos reportes sobre modelos estadísticos, análisis de varianza y comparación de medias para el DCL en experimentos factoriales (Tirado y Tirado, 2017; González *et al.*, 2019), particularmente cuando se utiliza un arreglo en PD o PSD con submuestreo balanceado y con la aplicación de algún paquete estadístico (Gomez y Gomez, 1984; Martínez, 1994; Montgomery, 2009; PROC ANOVA: Latin Square Split Plot :: SAS/STAT(R) 9.22 User's Guide). En el contexto previo, en esta investigación se construyó un modelo estadístico y se generaron las fórmulas para calcular GL y SC con dos metodologías, como un prerrequisito para su validación con algún software, como InfoStat, InfoGen o SAS, entre otros.

III. MATERIALES Y MÉTODOS

3.1 Simbología utilizada

En esta investigación se utilizará la misma terminología que emplearon Mendenhall (1987), Sahagún (2007 a), Montgomery (2009), Pérez *et al.* (2022) y González *et al.* (2023, 2004 a, b). Los cinco factores de clasificación que serán considerados en la serie de experimentos sin submuestreo balanceado o para cada experimento, sin y con submuestreo balanceado, estarán relacionados con Ambientes, Hileras, Columnas, A, B, que en lo sucesivo serán identificados como E, H, C, A y B, respectivamente; los subíndices y sus niveles también corresponderán a: $m= 1, 2, 3, \dots, e$; $i=1,2,3,\dots,h$; $j=1,2,3,\dots,c$; $k=1,2,3,\dots,t$; $l=1,2,3,\dots,b$.

3.2 Modelo estadístico

El modelo estadístico y las fórmulas para calcular grados de libertad y suma de cuadrados en un DCL en PD para un solo ensayo fue propuesto por Rodríguez *et al.* (2025) y, en la situación más práctica, su serie de experimentos se puede construir suponiendo ausencia de interacción entre Hileras (H) y Columnas (C), entre cada uno de éstas o de ambas con los factores E, A, B, y de H y C con las interacciones que se pueden generar con estos tres últimos. Con cinco factores de clasificación, E, H, C, A y B, habrá un total de 32 interacciones posibles, que se pueden obtener aplicando combinaciones de 5 factores tomados de cero, uno, dos, tres, cuatro y cinco veces a la vez ($C_0^5 + C_1^5 + C_2^5 + C_3^5 + C_4^5 + C_5^5$), pero de éstas no son viables las dos primeras debido a que es imposible formar combinaciones entre ninguno de ellos o entre las correspondientes a un solo factor. Por lo tanto, sólo son posibles: $:(\frac{1}{120}) n(n-1)\{\{60 + (n-2)\}\{20 + (n-3)[5 + (n-4)]\}\} = (\frac{1}{120}) 5(4)\{\{60 + 3\{20 + 2(5 + 1)\}\}\} = (\frac{1}{120}) 20(60 + 96) = 26$ interacciones, pero este número es más fácil calcularlo cómo: $32 - 1 - 6 = 26$

(interacciones de primero, segundo, tercero y cuarto orden). Con este antecedente es más fácil construir el modelo de referencia, particularmente si es aplicada correctamente la guía publicada por Sahagún (1998), quien, antes de introducir el principio de anidamiento, sugiere construir un modelo preliminar suponiendo que dichos factores solo tienen relaciones de cruzamiento. Así, el modelo propuesto es:

$$Y_{ijklm} = \mu + E_m + H_i + C_j + A_k + (EA)_{mk} + (EHA)_{i(mk)} + B_l + (EB)_{ml} + (AB)_{kl} + (EAB)_{mkl} + \epsilon_{ijklm}$$

Dónde: Y es la variable que será analizada; μ es la media aritmética general; E_m , H_i , C_j , A_k , B_l , son los efectos principales causados por los factores identificados como Ambientes, Hileras, Columnas, A, B, respectivamente; $(EHA)_{i(mk)}$ es el Error a; $(AB)_{kl}$, $(EA)_{mk}$, $(EB)_{ml}$, $(EAB)_{mkl}$ son interacciones de primer y segundo orden originadas por la combinación de dos o tres de los factores previamente descritos; ϵ_{ijklm} es el Error b o residual del modelo.

3.3 Diseño experimental

Para conformar una serie de experimentos en PD en DCL, con años y localidades confundidos en ambientes, para cada ensayo los niveles del Factor A son distribuidos en las Parcelas Principales con base en un diseño en Cuadro Latino, como lo sugirieron Smith (1951), Martínez (1994) y Tirado y Tirado (2017), y los niveles del Factor B se aleatorizarán al azar de manera independiente en cada subparcela.

3.4 Software recomendable

Para obtener un plan de aleatorización para cada experimento en este tipo de diseño experimental se puede utilizar el paquete estadístico STAR (Statistical Tools for Agricultural Research), del Instituto Internacional de Investigación en Arroz (IRRI), con sede en los Baños,

Filipinas (Digital Tools |International Rice Research Institute (irri.org). También se puede emplear SAS (Statistical Analysis System; SAS OnDemand for Academics | SAS) u otros paquetes estadísticos que dispongan de este tipo de herramientas para generarlos. Para realizar cálculos manuales con formas cuadráticas o matriciales podría utilizarse la calculadora científica disponible gratuitamente en <https://matrixcalc.org/es/>. Para analizar datos podrían utilizarse los paquetes estadísticos referenciados previamente u otros como InfoGen y InfoStat (Balzarini *et al.*, 2008; Di Rienzo *et al.*, 2008; Balzarini y Di Rienzo, 2016), entre otros.

3.5 Resultados

En Parcelas Principales, el número de niveles para Hileras (H), Columnas (C) y Factor A es igual ($h=c=t$); la combinación ht ó ct es equivalente a t^2 . Para un DBCA en PD, H ó C, pero no ambos, podría considerarse como r , que sería el número de repeticiones por cada combinación AB elegido para el experimento. La serie de experimentos se genera agregando E, el número de ambientes por evaluar. Los resultados que se muestran en la presente investigación son una extensión del caso propuesto por Rodríguez *et al.* (2023), quienes analizaron datos provenientes de experimentos monofactoriales conducidos en los diseños experimentales completamente al azar, bloques completos al azar y cuadro latino, sin y con submuestreo balanceado empleando InfoStat, InfoGen y SAS y, representan un enfoque complementario a la propuesta metodológica realizada por Zamudio y Alvarado (1996), con aplicación de SAS.

IV. RESULTADOS Y DISCUSIÓN

4.1 Primer artículo

Fórmulas para serie de experimentos en cuadro latino en arreglo de parcelas divididas

Artículo que será publicado en la Revista Mexicana de Ciencias Agrícolas, Volumen 16(6),
2025.

Fórmulas para serie de experimentos en cuadro latino en arreglo de parcelas divididas

Formulas for series of experiments in Latin Square Design in Split plot arrangement

Delfina de Jesús Pérez López, Artemio Balbuena Melgarejo, Jesús Hernández Ávila, J. Ramón Pascual Franco Martínez, José Antonio Rodríguez González², y Andrés González Huerta^{1δ}

¹Universidad Autónoma del Estado de México (UAEMéx), Facultad de Ciencias Agrícolas (FCAgr.), Centro de Investigación y Estudios Avanzados en Fitomejoramiento (C.I.E.A.F.). Campus Universitario “El Cerrillo”. El Cerrillo Piedras Blancas, Municipio de Toluca de Lerdo, Estado de México, México (CUC-CPB-MTEM). Apartado Postal 435. Tel. + 52 722 29 655 18. Ext. 60382. ² Programa de Maestría en Ciencias Agropecuarias y Recursos Naturales, UAEMéx-FCAgr.- CUC-CPB-MTEM. Tel. + 52 (722) 29 655 52. Ext. 117 (djperzl@uaemex.mx; jhernandez@uaemex.mx; jrfrancom@uaemex.mx; abalbuenam@uaemex.mx; jrodriguezg008@alumno.uaemex.mx). □ Autor para correspondencia: agonzalezh@uaemex.mx

Resumen

Las series de experimentos en tiempo, en espacio, o combinando ambos ambientes en arreglo de Parcelas Divididas o Subdivididas en Cuadro Latino no han sido utilizadas frecuentemente. En esta investigación se construye un modelo estadístico, y las fórmulas para obtener grados de libertad (GL) y suma de cuadrados (SC) usando formas cuadráticas o matriciales y mínimos cuadrados, cuando las componentes de tiempo y espacio están confundidas en ambientes, como un prerrequisito para extender su análisis con submuestreo balanceado. Con relación a la Parcela Principal se parte de la suposición de ausencia de interacción entre Hileras (H), Columnas (C), y niveles del Factor A, y de H y C con los ambientes; esta restricción también aplica entre H y C, o ambos, con los Factores A, B y su interacción, pero también se indica cómo llegar a los

mismos resultados introduciendo el principio de cruzamiento y anidamiento, particularmente si será aplicado algún paquete estadístico; adicionalmente se hace énfasis en las fórmulas para calcular directamente GL y SC para los Errores a y b, así como para los correspondientes a Parcelas Principales y Subparcelas.

Palabras clave: Modelo estadístico, años y localidades confundidos en ambientes, matrices.

Abstract

Series of experiments in time, space, or combining both environments in the arrangement of Split or Split - Split Plot in Latin Square Design have not been frequently used. In this research, a statistical model and formulas are constructed to obtain degrees of freedom (GL) and sum of squares (SC) using quadratic or matrix forms and least squares, when the components of time and space are confused in environments, as a prerequisite to extend their analysis with balanced subsampling. In relation to the Main Plot, it is assumed that there is no interaction between Rows (H), Columns (C), and levels of Factor A, and H and C with the environments; this restriction also applies between H and C, or both, with Factors A, B and their interaction, but it is also indicated how to arrive at the same results by introducing the principle of crossover and nesting, particularly if a statistical package will be applied; additionally, emphasis is placed on the formulas to directly calculate GL and SC for Errors a and b, as well as for those corresponding to Main Plots and Subplots.

Keywords: Statistical model, years and localities confused in environments, matrices.

Introducción

El análisis y la discusión de las series de experimentos en las ciencias agropecuarias y forestales, entre otras disciplinas, han sido abordadas frecuentemente desde la construcción de sus modelos

estadístico-genético, la obtención de un análisis de varianza y la aplicación de una metodología para la comparación de medias entre niveles de dos o más factores (Sahagún, 1993; 1994; 2007). El procesamiento estadístico de datos provenientes de arreglos en Parcelas Divididas (PD) o Subdivididas (PSD) se puede realizar de manera fácil y confiable utilizando InfoStat y InfoGen (Balzarini *et al.*, 2008; Di Rienzo *et al.*, 2008; Balzarini y Di Rienzo, 2016) y SAS, así como el correspondiente a las series de experimentos a través de años, localidades o ambos (González *et al.*, 2019; Pérez *et al.*, 2022; González *et al.*, 2024 b). Los diseños experimentales que más se han utilizado son Completamente al Azar (DCA) y Bloques Completos al Azar (DBCA) (Martínez, 1988; Gomez y Gomez, 1984; Little y Hills, 2008; Montgomery, 2009), pero Ledolter (2010) analizó y discutió los arreglos en PD para ensayos factoriales y para factoriales fraccionados, con énfasis en los diseños experimentales DCA y DBCA.

La serie de experimentos para un DCL, en PD o PSD podría ser fundamental para ayudar a detectar diferencias importantes entre años, entre localidades o entre ambos cuando se elijan diversos paquetes tecnológicos o diferentes manejos agronómicos, orientados a la generación, validación, aplicación y/o transferencia de tecnología a terrenos de agricultores, con el propósito de incrementar la productividad en los cultivos, y para mejorar la calidad de la materia prima que se deriva de ellos y que será utilizada en la agroindustria mexicana. El mejoramiento genético, la producción de semillas o la hibridación intra e interespecífica también podría documentarse mejor con base en estas metodologías estadísticas (González *et al.*, 2019; Pérez *et al.*, 2022; González *et al.*, 2024 a, b). Los terrenos de agricultores que son muy heterogéneos también presentan este tipo de variabilidad aleatoria que es indeseable para establecer un experimento adecuado; dicha heterogeneidad ambiental podría controlarse más eficientemente

utilizado un PD o un PSD en DCL (González *et al.*, 2019; Pérez *et al.*, 2022; Rodríguez *et al.*, 2025).

También se han encontrado pocos reportes sobre modelos estadísticos, análisis de varianza y comparación de medias para el DCL en experimentos factoriales (Tirado y Tirado, 2017; González *et al.*, 2019), particularmente cuando se utiliza un arreglo en PD o PSD con submuestreo balanceado y con la aplicación de algún paquete estadístico (Gomez y Gomez, 1984; Martínez, 1994; Montgomery, 2009; PROC ANOVA: Latin Square Split Plot :: SAS/STAT(R) 9.22 User's Guide). En el contexto previo, en esta investigación se construyó un modelo estadístico y se generaron las fórmulas para calcular GL y SC con dos metodologías, como un prerrequisito para su validación con algún software, como InfoStat, InfoGen o SAS, entre otros.

Materiales y métodos

Simbología

En esta investigación se utilizará la misma terminología que emplearon Mendenhall (1987), Sahagún (2007 a), Montgomery (2009), Pérez *et al.* (2022) y González *et al.* (2023, 2004 a, b). Los cinco factores de clasificación que ahora serán considerados son Ambientes, Hileras, Columnas, A, B, que en lo sucesivo serán identificados como E, H, C, A y B, respectivamente; los subíndices y sus niveles también corresponderán a: $m= 1, 2, 3, \dots, e$; $i=1,2,3,\dots,h$; $j=1,2,3,\dots,c$; $k=1,2,3,\dots,t$; $l=1,2,3,\dots,b$.

Modelo estadístico

El modelo estadístico y las fórmulas para calcular grados de libertad y suma de cuadrados en un DCL en PD para un solo ensayo fue propuesto por Rodríguez *et al.* (2025) y, en la situación más

práctica, su serie de experimentos se puede construir suponiendo ausencia de interacción entre Hileras (H) y Columnas (C), entre cada uno de éstas o de ambas con los factores E, A, B, y de H y C con las interacciones que se pueden generar con estos tres últimos. Con cinco factores de clasificación, E, H, C, A y B, habrá un total de 32 interacciones posibles, que se pueden obtener aplicando combinaciones de 5 factores tomados de cero, uno, dos, tres, cuatro y cinco veces a la vez ($C_0^5 + C_1^5 + C_2^5 + C_3^5 + C_4^5 + C_5^5$), pero de éstas no son viables las dos primeras debido a que es imposible formar combinaciones entre ninguno de ellos o entre las correspondientes a un solo factor. Por lo tanto, sólo son posibles: $(\frac{1}{120}) n(n-1) \{ \{ 60 + (n-2) \{ 20 + (n-3) [5 + (n-4)] \} \} \} = (\frac{1}{120}) 5(4) \{ \{ 60 + 3 \{ 20 + 2(5 + 1) \} \} \} = (\frac{1}{120}) 20(60 + 96) = 26$ interacciones, pero este número es más fácil calcularlo cómo: $32 - 1 - 6 = 26$ (interacciones de primero, segundo, tercero y cuarto orden). Con este antecedente es más fácil construir el modelo de referencia, particularmente si es aplicada correctamente la guía publicada por Sahagún (1998), quien, antes de introducir el principio de anidamiento, sugiere construir un modelo preliminar suponiendo que dichos factores solo tienen relaciones de cruzamiento. Así, el modelo propuesto es:

$$Y_{ijklm} = \mu + E_m + H_i + C_j + A_k + (EA)_{mk} + (EHA)_{i(mk)} + B_l + (EB)_{ml} + (AB)_{kl} + (EAB)_{mkl} + \varepsilon_{ijklm}$$

Dónde: Y es la variable que será analizada; μ es la media aritmética general; E_m , H_i , C_j , A_k , B_l , son los efectos principales causados por los factores identificados como Ambientes, Hileras, Columnas, A, B, respectivamente; $(EHA)_{i(mk)}$ es el Error a; $(AB)_{kl}$, $(EA)_{mk}$, $(EB)_{ml}$, $(EAB)_{mkl}$ son interacciones de primer y segundo orden originadas por la combinación de dos o tres de los factores previamente descritos; ε_{ijklm} es el Error b o residual del modelo.

Diseño experimental

Para conformar una serie de experimentos en PD en DCL, con años y localidades confundidos en ambientes, para cada ensayo los niveles del Factor A son distribuidos en las Parcelas Principales con base en un diseño en Cuadro Latino, como lo sugirieron Smith (1951), Martínez (1994) y Tirado y Tirado (2017), y los niveles del Factor B se aleatorizarán al azar de manera independiente en cada subparcela.

Software recomendable

Para obtener un plan de aleatorización para cada experimento en este tipo de diseño experimental se puede utilizar el paquete estadístico STAR (Statistical Tools for Agricultural Research), del Instituto Internacional de Investigación en Arroz (IRRI), con sede en los Baños, Filipinas (Digital Tools | International Rice Research Institute (irri.org)). También se puede emplear SAS (Statistical Analysis System; SAS OnDemand for Academics | SAS) u otros paquetes estadísticos que dispongan de este tipo de herramientas para generarlos. Para realizar cálculos manuales con formas cuadráticas o matriciales podría utilizarse la calculadora científica disponible gratuitamente en <https://matrixcalc.org/es/>. Para analizar datos podrían utilizarse los paquetes estadísticos referenciados previamente u otros como InfoGen y InfoStat (Balzarini *et al.*, 2008; Di Rienzo *et al.*, 2008; Balzarini y Di Rienzo, 2016), entre otros.

Resultados

En Parcelas Principales, el número de niveles para Hileras (H), Columnas (C) y Factor A es igual ($h=c=t$); la combinación ht ó ct es equivalente a t^2 . Para un DBCA en PD, H ó C, pero no ambos, podría considerarse como r , que sería el número de repeticiones por cada combinación AB elegido para el experimento. La serie de experimentos se genera agregando E, el número de

ambientes por evaluar. Los resultados que a continuación se presentan son una extensión del caso propuesto por Rodríguez *et al.* (2025).

Fórmulas para Grados de Libertad (GL)

$$GL \text{ Total} = et^2b - 1$$

$$GL \text{ E} = e - 1$$

$$GL \text{ H} = h - 1 = t - 1; GL \text{ C} = c - 1 = t - 1; GL \text{ A} = t - 1; GL \text{ ExA} = (e - 1)(t - 1);$$

$$GL \text{ Error a} = t [e(t - 1) - 2] + 2$$

$$GL \text{ Parcelas Principales (PP)} = GL \text{ E} + GL \text{ H} + GL \text{ C} + GL \text{ A} + GL \text{ ExA} + GL \text{ Error a}$$

$$\text{Para verificación: } GL \text{ PP} = et^2 - 1$$

$$GL \text{ B} = b - 1; GL \text{ AxB} = (t - 1)(b - 1)$$

$$GL \text{ ExB} = (e - 1)(b - 1)$$

$$GL \text{ ExAxB} = (e - 1)(t - 1)(b - 1)$$

$$GL \text{ Error b} = et(t - 1)(b - 1)$$

$$GL \text{ Subparcelas (SUB)} = GL \text{ Total} - GL \text{ E} - GL \text{ H} - GL \text{ C} - GL \text{ A} - GL \text{ ExA} - GL \text{ Error a} -$$

$$GL \text{ B} - GL \text{ AxB} - GL \text{ ExB} - GL \text{ ExAxB} - GL \text{ Error b}$$

$$\text{También: } GL \text{ SUB} = GL \text{ B} + GL \text{ AxB} + GL \text{ ExB} + GL \text{ ExAxB} + GL \text{ Error b}$$

$$\text{Para verificación: } GL \text{ SUB} = et^2 (b - 1)$$

$$GL \text{ PP} + GL \text{ SUB} = (et^2 - 1) + et^2(b - 1) = et^2b - 1 = GL \text{ Total}$$

Fórmulas para Suma de Cuadrados (SC)

En el denominador de las siguientes fórmulas, h ó c es nulo; E, H, C, A, B estarán relacionados con los subíndices m, i, j, k, l, respectivamente. Las formas cuadráticas o matriciales se escribirán como lo hicieron González *et al.* (2023), y González *et al.* (2024 a, b); en éstas, antes de su aplicación, se utilizarán sumas o totales realizadas sobre el o los subíndices que no se muestran en su numerador; la matriz J es cuadrada y sólo está formada por unos y tiene t^2b hileras y columnas.

$$SC\ Tot = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^e Y_{ijklm}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^b \sum_{l=1}^e Y_{ijklm}}{t^2be} \right)^2 = Y'Y - \left(\frac{1}{t^2be} \right)$$

$Y'JY$

$$SC\ H = \left(\frac{1}{tbe} \right) \sum_{i=1}^h Y_{i\dots}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^b \sum_{l=1}^e Y_{ijklm}}{t^2be} \right)^2 = \left(\frac{1}{tbe} \right) Y'_{i\dots} Y_{i\dots} - \left(\frac{1}{t^2be} \right) Y'JY$$

$$SC\ C = \left(\frac{1}{tbe} \right) \sum_{j=1}^c Y_{.j\dots}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^b \sum_{l=1}^e Y_{ijklm}}{t^2be} \right)^2 = \left(\frac{1}{tbe} \right) Y'_{.j\dots} Y_{.j\dots} - \left(\frac{1}{t^2be} \right) Y'JY$$

$$SC\ A = \left(\frac{1}{tbe} \right) \sum_{k=1}^t Y_{..k\dots}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^b \sum_{l=1}^e Y_{ijklm}}{t^2be} \right)^2 = \left(\frac{1}{tbe} \right) Y'_{..k\dots} Y_{..k\dots} - \left(\frac{1}{t^2be} \right) Y'JY$$

$$SC\ E = \left(\frac{1}{t^2b} \right) \sum_{m=1}^e Y_{\dots m}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^b \sum_{l=1}^e Y_{ijklm}}{t^2be} \right)^2 = \left(\frac{1}{t^2b} \right) Y'_{\dots m} Y_{\dots m} - \left(\frac{1}{t^2be} \right) Y'JY$$

$SC\ TRAT1 = SC\ E + SC\ A + SC\ ExA$. Así: $SC\ ExA = SC\ TRAT1 - SC\ E - SC\ A$

$$\text{Adicionalmente: } SC\ TRAT1 = \left(\frac{1}{tb} \right) \sum_{k=1}^t \sum_{m=1}^e Y_{..k.m}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^b \sum_{l=1}^e Y_{ijklm}}{t^2be} \right)^2$$

$$= \left(\frac{1}{tb} \right) Y'_{..k.m} Y_{..k.m} - \left(\frac{1}{t^2be} \right) Y'JY$$

$$SC\ Parcelas\ Principales\ (PP) = \left(\frac{1}{b} \right) \sum_{i=1}^h \sum_{k=1}^t \sum_{m=1}^e Y_{i.k.m}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^b \sum_{l=1}^e Y_{ijklm}}{t^2be} \right)^2$$

$$= \left(\frac{1}{b} \right) Y'_{i.k.m} Y_{i.k.m} - \left(\frac{1}{t^2be} \right) Y'JY$$

$$\text{SC Error a} = \left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t \sum_{m=1}^e Y_{i.k.m}^2 - \left(\frac{1}{tbe}\right) \sum_{i=1}^h Y_{i....}^2 -$$

$$\left(\frac{1}{tbe}\right) \sum_{j=1}^c Y_{j...}^2 - \left(\frac{1}{tb}\right) \sum_{k=1}^t \sum_{m=1}^e Y_{..k.m}^2 + 2 \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^e Y_{ijklm}}{t^2 be}\right)^2$$

También:

$$\text{SC Error a} = \left(\frac{1}{b}\right) Y'_{i.k.m} Y_{i.k.m} - \left(\frac{1}{tbe}\right) Y'_{i....} Y_{i....} - \left(\frac{1}{tbe}\right) Y'_{j...} Y_{j...} - \left(\frac{1}{tb}\right) Y'_{..k.m} Y_{..k.m} + \left(\frac{2}{t^2 be}\right) Y' J Y$$

$$\text{SC B} = \left(\frac{1}{t^2 e}\right) \sum_{l=1}^b Y_{...l}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^e Y_{ijklm}}{t^2 be}\right)^2 = \left(\frac{1}{t^2 e}\right) Y'_{...l} Y_{...l} - \left(\frac{1}{t^2 be}\right) Y' J Y$$

Para calcular SC AxB, deberá obtenerse primero SC TRAT2 cómo:

$$\text{SC TRAT2} = \text{SC A} + \text{SC B} + \text{SC AxB}. \text{ Entonces: } \text{SC AxB} = \text{SC TRAT2} - \text{SC A} - \text{SC B}$$

$$\text{Dónde: } \text{SC TRAT2} = \left(\frac{1}{te}\right) \sum_{k=1}^t \sum_{l=1}^b Y_{..kl}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^e Y_{ijklm}}{t^2 be}\right)^2 = \left(\frac{1}{te}\right) Y'_{..kl} Y_{..kl} -$$

$$\left(\frac{1}{t^2 be}\right) Y' J Y$$

$$\text{Cómo } \text{SC TRAT3} = \text{SC E} + \text{SC B} + \text{SC ExB}. \text{ SC ExB} = \text{SC TRAT3} - \text{SC E} - \text{SC B}$$

$$\text{Pero, } \text{SC TRAT3} = \left(\frac{1}{t^2}\right) \sum_{l=1}^b \sum_{m=1}^e Y_{...lm}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^e Y_{ijklm}}{t^2 be}\right)^2 = \left(\frac{1}{t^2}\right) Y'_{...lm} Y_{...lm} -$$

$$\left(\frac{1}{t^2 be}\right) Y' J Y$$

$$\text{SC TRAT4} = \text{SC E} + \text{SC A} + \text{SC B} + \text{SC ExA} + \text{SC ExB} + \text{SC AxB} + \text{SC ExAxB}$$

$$\text{Así : } \text{SC ExAxB} = \text{SC TRAT4} - \text{SC E} - \text{SC A} - \text{SC B} - \text{SC ExA} - \text{SC ExB} - \text{SC AxB}$$

Para cálculo directo puede utilizarse la siguiente formula:

$$\text{SC ExAxB} = \left(\frac{1}{t}\right) \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^e Y_{..klm}^2 - \left(\frac{1}{tb}\right) \sum_{k=1}^t \sum_{m=1}^e Y_{..k.m}^2 - \left(\frac{1}{t^2}\right) \sum_{l=1}^b \sum_{m=1}^e Y_{...lm}^2 +$$

$$\left(\frac{1}{t^2 b}\right) \sum_{m=1}^e Y_{...m}^2 - \left(\frac{1}{te}\right) \sum_{k=1}^t \sum_{l=1}^b Y_{..kl}^2 + \left(\frac{1}{tbe}\right) \sum_{k=1}^t Y_{..k..}^2 + \left(\frac{1}{t^2 e}\right) \sum_{l=1}^b Y_{...l}^2 -$$

$$\left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^e Y_{ijklm}}{t^2 b e}\right)^2 = \left(\frac{1}{t}\right) Y'_{..klm} Y_{..klm} - \left(\frac{1}{tb}\right) \sum_{k=1}^t \sum_{m=1}^e Y_{..k.m} - \left(\frac{1}{t^2}\right) Y'_{...lm} Y_{...lm} +$$

$$\left(\frac{1}{t^2 b}\right) Y'_{...m} Y_{...m} - \left(\frac{1}{te}\right) Y'_{..kl} Y_{..kl} + \left(\frac{1}{tbe}\right) Y'_{..k..} Y_{..k..} + \left(\frac{1}{t^2 e}\right) Y'_{...l} Y_{...l} - \left(\frac{1}{t^2 be}\right) Y' Y Y$$

Finalmente:

$$\text{SC Error b} = \text{SC Total} - \text{SC E} - \text{SC H} - \text{SC C} - \text{SC A} - \text{SC ExA} - \text{SC Error a} - \text{SC B} - \text{SC}$$

$$\text{AxB} - \text{SC ExB} - \text{SC ExAxB}$$

También es válida:

$$\text{SC Error b} = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^e Y_{ijklm}^2 -$$

$$\left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t \sum_{m=1}^e Y_{i.k.m}^2 - \left(\frac{1}{t}\right) \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^e Y_{..klm}^2 + \left(\frac{1}{tb}\right) \sum_{k=1}^t \sum_{m=1}^e Y_{..k.m}$$

$$= Y' Y - \left(\frac{1}{b}\right) Y'_{i.k.m} Y_{i.k.m} - \left(\frac{1}{t}\right) Y'_{..klm} Y_{..klm} + \left(\frac{1}{tb}\right) Y'_{..k.m} Y_{..k.m}$$

SC Subparcelas (SUB) = SC Total – SC Parcelas Principales (PP)

$$\text{SC SUB} = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^e Y_{ijklm}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^e Y_{ijklm}}{t^2 b e}\right)^2 -$$

$$\left\{ \left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t \sum_{m=1}^e Y_{i.k.m}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^e Y_{ijklm}}{t^2 b e}\right)^2 \right\}$$

$$= \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^e Y_{ijklm}^2 - \left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t \sum_{m=1}^e Y_{i.k.m}^2 = Y' Y - \left(\frac{1}{b}\right) Y'_{i.k.m} Y_{i.k.m}$$

Discusión

Cuando se diseña y analiza un experimento o una serie de ensayos a través de años (A), localidades (L) o en combinaciones de éstos (AxL), están implícitas al menos tres etapas fundamentales: a) la construcción o elección de un modelo estadístico, b) la generación de un análisis de varianza y, c) la aplicación de una prueba de comparación de medias aritméticas de

dos niveles diferentes, en uno, dos o más factores bajo consideración (Sahagún, 1993; 1998; 2007). Con relación a esta estrategia también es recomendable la aplicación de uno o más paquetes estadísticos para ahorrar tiempo durante el análisis estadístico de los datos. En el contexto previo, la elección del software adecuado es también de gran relevancia para generar las salidas requeridas de manera fácil y confiable (González *et al.*, 2023; González *et al.*, 2024 a, b).

Las series de experimentos que han sido más frecuentemente abordadas en la investigación agronómica son las que corresponden a los diseños completamente al azar (DCA) y bloques completos al azar (DBCA), para los arreglos combinatorio, en parcelas divididas (PD), parcelas subdivididas (PSD), o bloques o franjas divididas (FD) (Sahagún, 1993, 1994, 2007; Gomez y Gomez, 1984; Ledolter, 2010; Tirado y Tirado, 2017), pero el DCL en PD, PSD o FD no está bien documentado.

En la literatura publicada se ha encontrado poca evidencia con relación al modelo estadístico de referencia, así como para un Análisis de varianza (Anava) y una comparación de medias (Ledolter, 2010; Tirado y Tirado, 2017; <https://biometrics.ilri.org/Publication/Full%20Text/chapter20.pdf>), especialmente para submuestreo (Gomez y Gomez, 1984; Martínez, 1988; PROC ANOVA: Latin Square Split Plot :: SAS/STAT® 9.22 User's Guide). Martínez (1994) sólo presentó una tabla con las fuentes de variación (FV) y los grados de libertad (GL) que podrían calcularse sin submuestreo. Gomez y Gomez (1984) mostraron dos tablas con las FV y los GL para un DBCA y para un arreglo en PD con submuestreo balanceado, pero para un solo ensayo; adicionalmente, ellos describieron los procedimientos para obtener un Anava en ambas situaciones, con énfasis en la estimación de los Errores Muestral (EM) y Experimental (EE), que en el presente estudio conforman al Error Conjunto (EC).

En el Sistema para Análisis Estadístico (SAS, Statistical Analysis System) sólo se presentaron los datos proporcionados por Smith (1951; PROC ANOVA: Latin Square Split Plot :: SAS/STAT® 9.22 User's Guide) pero sin submuestreo, así como un código para obtener un Anava con tres tipos de errores; con relación a la presente investigación y para un solo ensayo, $CxHxA$ es igual al Error a, y la suma de HxB y del residual del modelo originan al Error b. Tirado y Tirado (2017) presentaron el modelo estadístico para un PD en DCL sin submuestreo, así como un ejemplo para generar su Anava; ellos también generalizaron este tipo de análisis para un arreglo en Parcelas Subdivididas en DCL sin submuestreo.

En estudios previos, Sahagún (1993; 1994; 2007) abordó varias situaciones en las cuales analizó las series de ensayos para los diseños experimentales DCA y DBCA, y de éste último en arreglo en PD; él analiza sus modelos, las esperanzas de los cuadrados medios para realizar las pruebas de hipótesis pertinentes en un análisis de varianza, así como su eficiencia, con y sin restricción en dichos modelos. Adicionalmente, él consideró dos casos: a) cuando hay cruzamiento entre años, localidades y genotipos y, b) cuando los años están anidados dentro de las localidades, por lo que los genotipos pueden o no estar anidados en la interacción AxL .

Rodríguez *et al.* (2025) consideraron el modelo estadístico y las fórmulas para calcular grados de libertad y suma de cuadrados aplicando dos metodologías al caso específico de un ensayo conducido en un diseño experimental en cuadro latino en arreglo de parcelas divididas, como un prerrequisito para extender su análisis a serie de experimentos con la aplicación de software, cuando años y localidades están confundidos dentro de ambientes contrastantes, por lo que la presente investigación es una extensión de su trabajo, como un prerrequisito para analizar datos provenientes de este tipo de experimentos cuando aplica el submuestreo balanceado dentro de las unidades experimentales que conforman a cada ensayo. Sin embargo, aún quedará pendiente

el análisis y discusión de este tipo de ensayos cuando, como Sahagún (1993; 1994; 2007) lo sugirió, los años y las localidades sean consideradas como factores cruzados o cuando los años sean anidados dentro de localidades, particularmente para el caso de un PD, PSD y FD en DCL.

Una situación complementaria a lo considerado previamente fue abordada por González *et al.* (2024 b), quienes fraccionaron cada una de las repeticiones en arreglo de “g” grupos de bloques completos balanceados para un DBCA, confundiendo años y localidades en ambientes diferentes; ellos extendieron el caso presentado por Gomez y Gomez (1984), para un solo ensayo, para el análisis de 45 cultivares de arroz (*Oriza sativa* L.) clasificados en tres grupos, con 15 cultivares dentro de cada uno de éstos. Ellos también consideraron la posibilidad de agrupar el material genético con base en dos criterios: a) que los grupos fueran tan heterogéneos como sea posible, y b) que dentro de cada grupo existieran diferencias tan pequeñas en la madurez fisiológica en la fracción de material que está siendo considerada; si no hubiera diferencias significativas entre grupos de cultivares el análisis estadístico de los datos puede realizarse simplemente como un DBCA, o como una serie de ensayos en este tipo de diseño experimental.

En la serie de experimentos que está siendo discutida en el presente estudio, para un arreglo en PD en DCL sin submuestreo, se asume que el modelo estadístico fue construido considerando que años y localidades están confundidos en ambientes contrastantes (E), y éstos últimos no presentan interacción con Hileras (H), columnas ® o con ambos, pero E está cruzado con los factores A y B. Adicionalmente, H, C o ambos tampoco presentan interacción con A, B o AB. Finalmente, cuando se aplique InfoStat el Error a estará representado por las interacciones ExHxA o ExCxA, debido a la restricción impuesta en las parcelas principales ($h=c=t$),

relacionada con el doble bloqueo que se aplica a las unidades experimentales cuando se aleatorizan los niveles del Factor A.

Para una serie de experimentos en DBCA, Sahagún (1993) consideró tres situaciones con relación a las repeticiones: a) éstas anidadas dentro de localidades, b) éstas anidadas dentro de años y localidades, c) éstas anidadas dentro de un Factor C, correspondiente a las combinaciones entre A y L. En los tres casos él recomendó considerar a repeticiones como factor aleatorio, al igual que A, L o C. Con relación a genotipos, con propósitos de selección, él sugirió considerarlos como factor de efectos fijos, pero si el propósito es estimar componentes de varianza, éste deberá definirse como factor aleatorio.

En el contexto previo, aplicando la guía propuesta por Sahagún (1998) se construirá el mismo modelo si se establecen las siguientes condiciones: a) Hileras (H) anidado dentro de Columnas ® o viceversa; H, C o ambos anidados dentro de ambientes ®, b) E, A y B o, sus interacciones, están cruzados, c) H, C o ambos están anidados dentro de B o AB, d) E es factor aleatorio, pero A y B pueden ser factores fijos, e) el residual del modelo está anidado en todas las componentes del modelo. Además de lo anterior, deberá considerarse si hay o no restricción en las componentes de dicho modelo (Sahagún, 1993).

Para propósitos complementarios deberán considerarse las publicaciones de Sahagún (1994 y 2007 a,b), quien propone probar las hipótesis estadísticas correctas relacionadas con las componentes de los modelos estadísticos que el analizó y discutió, utilizando los cuadrados medios apropiados. Esta última situación también es un punto crítico en el tema que está siendo considerado en la presente investigación o en otras que ahora están siendo utilizadas para generar, validar, aplicar o transferir tecnología, debido a que cuando se emplea un paquete estadístico podrían generarse análisis de varianza con valores de F incorrectos, así como para

alguna prueba de comparación de medias para efectos principales o interacciones que podrían no ser válidas también, como lo sugirieron Sahagún (1993; 1994 a, b; 1998; 2007), Gomez y Gomez (1984), y Montgomery (2009), entre otros.

Las metodologías presentadas en esta investigación permitirán homologar fórmulas para calcular grados de libertad y suma de cuadrados de manera fácil y confiable, como lo han mostrado para otros estudios González *et al.* (2023), y González *et al.* (2024 a, b); con su simbología y la descrita en Mendenhall (1987), Sahagún (2007 a), y Montgomery (2009), entre otros, se generarán de manera fácil y confiable las formas cuadráticas o matriciales que alimentarán más frecuentemente a las calculadoras de matrices, o a SAS, Agrobases, SPSS, StatGraphics, STAR, PB Tools, entre otros.

Conclusiones

En el modelo estadístico que se propone se asumió que años y localidades están confundidos en ambientes contrastantes \otimes , y que éstos últimos no presentan interacción con Hileras (H), columnas \otimes o con ambos, pero E está cruzado con los factores A y B. Adicionalmente, se consideró que H, C o ambos tampoco presentan interacción con A, B o AB. Finalmente, el Error a estará representado por las interacciones $E \times H \times A$ o $E \times C \times A$. Para verificar grados de libertad y suma de cuadrados en ambos modelos pueden aplicarse las fórmulas alternativas que fueron construidas para los Errores a y b, así como las correspondientes a Parcelas Principales y Subparcelas, para cada una o para ambas metodologías.

Literatura citada

Balzarini, M. G. y Di Rienzo, J. A. 2016. InfoGen. FCA. Universidad Nacional de Córdoba, Argentina (<http://www.info-Gen.com.ar>).

Balzarini, M. G.; González, L.; Tablada, M.; Casanoves, F.; Di Rienzo, J. A. y Robledo, C. W. 2008. Manual del usuario de InfoStat. Ed. Brujas, Córdoba, Argentina. 348 pp.

Di Rienzo, J. A.; Casanoves, F.; Balzarini, M. G.; González, L.; Tablada, M.; Robledo, C. W. 2008. InfoStat, versión 2008. Grupo InfoStat, FCA. Universidad Nacional de Córdoba, Argentina. (<https://www.infostat.com.ar>).

Gomez, K. A. y Gomez, A. A. 1984. Statistical procedures for agricultural research. 2nd Ed. John Wiley & Sons, Inc. Printed in Singapore. 680 pp.

González, H. A.; Pérez, L. D. J.; Rubí, A. M.; Gutiérrez, R. F.; Franco, M. J. R. P. y Padilla, L. A. 2019. InfoStat, InfoGen y SAS para contrastes mutuamente ortogonales en experimentos en bloques completos al azar en parcelas subdivididas. *Revista Mexicana de Ciencias Agrícolas*. 10(6):1417-1431.

González, H. A.; Pérez, L. D. J.; Balbuena, M. A.; Franco, M. J. R.; Gutiérrez, R. F. y Rodríguez, G. J. A. 2023. Submuestreo balanceado en experimentos monofactoriales usando InfoStat y InfoGen: validación con SAS. *Revista Mexicana de Ciencias Agrícolas*. 14(2):235-249 (DOI:<https://doi.org/10.29312/remexca.V14i2.3418>).

González, H., A.; Pérez, L., D.J.; Hernández, A., J.; Franco, M., J.R.P.; Rubí, A., M.; Balbuena, M., A. 2024 a. Tratamientos anidados dentro de un arreglo en grupos de bloques completos balanceados. *Revista Mexicana de Ciencias Agrícolas*.15(2), e 3634 (DOI: <https://doi.org/10.29312/remexca.v15i2.3634>).

González, H., A.; Pérez, L., D.J.; Hernández, A., J.; Franco, M., J.R.P.; Balbuena, M., A.; Rubí, A., M. 2024 b. Serie de experimentos para tratamientos anidados en grupos en arreglo de bloques completos balanceados. *Revista Mexicana de Ciencias Agrícolas*. 15(7) (En edición).

Ledolter, J. 2010. Split-plot design: 48discusión and examples. *International Journal of Quality Engineering and Technology*. 1(4): 441-457 (DOI: <https://doi.org/10.1504/IJQET.2010.035588>).

Little, T. M. y Hills, F. J. 2008. *Métodos Estadísticos para la Investigación en la Agricultura*. Ed. Trillas, SA de CV. México, DF. 270 p. ISBN:978-968-24-3629-1.

Martínez, G. A. 1988. *Diseños experimentales. Métodos y elementos de teoría*. Editorial. Trillas, 1ª Ed. México, DF. 756 pp.

Martínez, G., A. 1994. *Experimentación Agrícola: Métodos Estadísticos*. Primera Edición. Universidad Autónoma Chapingo, México, D.F., 357 p (ISBN: 9789 6888 42 867).

Mendenhall, W. 1987. *Introducción a la Probabilidad y la Estadística*. Grupo Editorial Iberoamérica. 1ª Ed. México, DF. 626 pp.

Montgomery, D. C. 2009. *Design and Analysis of Experiments*. 7th Ed. John Wiley & Sons, Inc. U.S.A. 656 pp.

Pérez, L., D.; Jasso, B., G.; Saavedra, G., C.; Franco, M., J.R.P.; Ramírez, D., J. F.; González, H., A. 2022. Uso de artificios en Opstat para analizar series de experimentos en dialélico parcial. *Revista Mexicana de Ciencias Agrícolas*. 13(2) 273-287 (DOI: <https://doi.org/10.29312/remexca.v13i2.3130>).

Rodríguez, G., J. A.; Pérez, L., D.J.; Hernández, A., J.; Balbuena, M., A.; Franco, M., J. R. P.; González, H., A. 2025. Parcelas Divididas en Cuadro Latino: Modelos estadísticos y fórmulas, sin y con submuestreo. *Revista Mexicana de Ciencias Agrícolas* (en revisión: julio del 2024).

Sahagún, C., J. 1993. Funcionalidad de cuatro modelos para las evaluaciones genotípicas en series de experimentos. *Revista Fitotecnia Mexicana* 16(3): 161-171.

Sahagún, C., J. 1994. Evaluación de genotipos en series de experimentos: diferencias en parámetros genéticos generados en dos modelos. *Revista Fitotecnia Mexicana* 17(2):116-125.

Sahagún, C., J. 1998. Construcción y análisis de los modelos fijos, aleatorios y mixtos. Departamento de Fitotecnia. Programa Nacional de Investigación en Olericultura. Universidad Autónoma Chapingo. Boletín Técnico Núm. 2, 64 p.

Sahagún, C., J. 2007 b. Evaluación de genotipos en heterogeneidad meteorológica intrarregional: Confusión vs. Anidamiento de años en localidades. *Revista Fitotecnia Mexicana* 30(1):97-104.

Sahagún, C. J. 2007^a. *Estadística Descriptiva y Probabilidad: Una Perspectiva Biológica*. 2^a Ed. Universidad Autónoma Chapingo (UACH). México, D.F. 282 pp. ISBN: 978 -968-02- 0357-4.

Smith, W., G. 1951. Dissertation notes on Canadian Sugar Factories Ltd, Taber, Alberta, Canadá.

Tirado, E., G., y Tirado, G., D. N. 2017. *Tratado de Estadística Experimental*. Primera Edición. Editorial Centro de Estudios e Investigaciones para el Desarrollo Docente (CENID, AC), Guadalajara, Jalisco, México (ISBN: 978-607-8435-43-2).

4.2 Segundo artículo

Parcelas Divididas en Cuadro Latino: Modelos estadísticos y fórmulas, sin y con submuestreo

Este artículo será publicado en la Revista Mexicana de Ciencias Agrícolas. Volumen 16(2),
2025.

Parcelas Divididas en Cuadro Latino: Modelos estadísticos y fórmulas, sin y con submuestreo

José Antonio Rodríguez González¹, Delfina de Jesús Pérez López², Jesús Hernández Ávila², Artemio Balbuena Melgarejo², J. Ramón Pascual Franco Martínez², y Andrés González Huerta^{2δ}

¹Programa de Maestría en Ciencias Agropecuarias y Recursos Naturales, Universidad Autónoma del Estado de México (UAEMéx), Facultad de Ciencias Agrícolas (FCAgr), Campus Universitario “El Cerrillo”. El Cerrillo Piedras Blancas, Municipio de Toluca de Lerdo, Estado de México, México (CUC-CPB-MTLM). Tel. + 52 (722) 29 655 52. Ext. 117 (pcarn@uaemex.mx). ²Centro de Investigación y Estudios Avanzados en Fitomejoramiento (C.I.E.A.F.), FCAgr – UAEMéx. CUC-CPB-MTLM. Apartado Postal 435, Tel: +52 722 29 655 18, Ext. 148 (jrodriguez008@alumno.uaemex.mx; djperezl@uaemex.mx; jhernandez@uaemex.mx; jrfrancom@uaemex.mx; abalbuenam@uaemex.mx. □ Autor para correspondencia: agonzalez@uaemex.mx

Resumen

Se ha publicado poca información para un arreglo en Parcelas Divididas para un diseño en Cuadro Latino. En esta investigación se construyen sus modelos estadísticos, y las fórmulas que permitan obtener grados de libertad y suma de cuadrados considerando las metodologías de Mínimos Cuadrados y formas cuadráticas o matriciales, sin y con submuestreo balanceado. Se parte de la suposición de que no hay interacción entre Hileras, Columnas, y Factor A, pero se indica como llegar a los mismos resultados introduciendo el principio de cruzamiento y anidamiento, particularmente si será aplicado InfoStat o InfoGen; se sugiere utilizar la diferencia principal entre las salidas que generan ambos paquetes estadísticos para subdividir al error

conjunto en error muestral y error experimental, pero también se hace énfasis en las fórmulas que fueron derivadas para su cálculo directo.

Palabras clave: Mínimos Cuadrados, formas cuadráticas o matriciales, experimentos bifactoriales.

Abstract

Little information has been published for a split-plot arrangement for a Latin Square design. In this research, statistical models are built, and the formulas that allow obtaining degrees of freedom and sum of squares considering the methodologies of Least Squares and quadratic or matrix forms, without and with balanced subsampling, were obtained. It is assumed that there is no interaction between Rows, Columns, and Factor A, but it is indicated how to reach the same results by introducing the principle of crossing and nesting, particularly if InfoStat or InfoGen will be applied. It is suggested to use the main difference between the outputs generated by both statistical packages to subdivide the joint error into sampling error and experimental error, but emphasis is also placed on the formulas that were derived for its direct calculation.

Key words: Least Squares, quadratic or matrix forms, two-factor experiments

Introducción

Los arreglos de unidades experimentales en Parcelas Divididas (PD) o Subdivididas (PSD) han sido utilizados ampliamente para estudiar efectos o varianzas en ensayos bi o trifactoriales (Balzarini *et al.*, 2008; Di Rienzo *et al.*, 2008; Balzarini y Di Rienzo, 2016; Tirado y Tirado,

2017), o en series de experimentos cuando éstos se extienden a diversos ambientes, conformados por años, localidades o combinaciones de ambos (González *et al.*, 2019; Pérez *et al.*, 2022). Los planes de aleatorización más utilizados corresponden a los diseños experimentales Completamente al Azar (DCA) y Bloques Completos al Azar (DBCA) (Martínez, 1988; Gomez y Gomez, 1984; Little y Hills, 2008; Montgomery, 2009). Ledolter (2010) realizó una discusión acerca de los arreglos de PD para ensayos completos y para experimentos factoriales fraccionados, con énfasis en DCA y DBCA, pero sin considerar submuestreo balanceado.

El DCL, en PD o PSD, no ha sido muy utilizado, pero podría ser de gran utilidad para evaluar las diferencias que originan años, localidades o ambos con distanciamiento entre plantas, fórmulas de fertilización, abonos orgánicos, insecticidas, fungicidas, herbicidas, métodos de labranza, láminas de riego, y fechas de corte en forrajes, entre otros, cuando éstos se asignen a parcela principal, y en subparcelas podrían considerarse especies vegetales o variedades contrastantes de alguna de éstas (González *et al.*, 2019; Pérez *et al.*, 2022). Los terrenos de agricultores que son muy heterogéneos también presentan este tipo de variabilidad aleatoria que es indeseable para establecer un experimento adecuado; ésta podría controlarse más eficientemente utilizando un PD en DCL u otro diseño experimental (González *et al.*, 2019; Pérez *et al.*, 2022).

También se han encontrado pocos reportes sobre los modelos estadísticos, el análisis de varianza y la comparación de medias (Tirado y Tirado, 2017; González *et al.*, 2019), particularmente cuando se aplica submuestreo balanceado y algún paquete estadístico (Gomez y Gomez, 1984; Martínez, 1994; Montgomery, 2009; PROC ANOVA: Latin Square Split Plot :: SAS/STAT® 9.22 User's Guide). Así, el objetivo principal de esta investigación fue construir sus modelos sin y con submuestreo balanceado, y presentar las fórmulas para calcular grados de libertad y

suma de cuadrados con base en dos metodologías, como un prerrequisito para la aplicación de algún paquete estadístico.

Materiales y métodos

Simbología

En este estudio la variable cuantitativa de interés será identificada con Y. Adicionalmente, será empleada la simbología descrita por Mendenhall (1987), Sahagún (2007) y Montgomery (2009). También será aplicada la terminología descrita por Pérez *et al.* (2022) y por González *et al.* (2023, 2004 a, b). Los factores de clasificación serán Hileras, Columnas, A, B, cuyos niveles serán, respectivamente: $i=1,2,3,\dots,h$; $j=1,2,3,\dots,c$; $k=1,2,3,\dots,t$; $l=1,2,3,\dots,b$. Con submuestreo balanceado, además de lo anterior, será incluido el Factor S ($m=1,2,3,\dots,s$).

Modelos estadísticos balanceados

Tirado y Tirado (2017) presentaron un modelo para PD en DCL sin submuestreo, pero los dos que se presentan a continuación se construyeron aplicando la guía publicada por Sahagún (1998). Ambos modelos también se pueden construir bajo la suposición de ausencia de interacción entre Hileras, Columnas y niveles del Factor A, como lo sugirieron Gomez y Gomez (1984), Martínez (1994), y Montgomery (2009).

Sin submuestreo:

$$Y_{ijkl} = \mu + H_i + C_j + A_k + (HA)_{ik} + B_l + (AB)_{kl} + \varepsilon_{ijkl}$$

Con submuestreo:

$$Y_{ijklm} = \mu + H_i + C_j + A_k + (HA)_{ik} + B_l + (AB)_{kl} + \delta_{m(ikl)} + \varepsilon_{ijklm}$$

Dónde: Y es la variable de interés; μ es la media aritmética general; H_i , C_j , A_k , B_l , son los efectos causados por Hileras, Columnas, y Factores A y B, respectivamente; $(HA)_{ik}$ es el Error a; $(AB)_{kl}$ es la interacción entre ambos factores; $\delta_{m(ikl)}$ es el error muestral; ε_{ijkl} y ε_{ijklm} son el error experimental, sin y con submuestreo.

Diseño experimental

Los niveles del Factor A serán asignados a Parcelas Principales con base en un diseño en Cuadro Latino, como lo sugirieron Smith (1951), Martínez (1994) y Tirado y Tirado (2017), y los niveles del Factor B se aleatorizan en Subparcelas de manera completamente al azar; en un modelo no se considerará submuestreo balanceado y en otro sí.

Software recomendable

Para obtener un plan de aleatorización para este tipo de experimento se puede utilizar el paquete estadístico STAR, del Instituto Internacional de Investigación en Arroz (IRRI), con sede en los Baños, Filipinas (Digital Tools | International Rice Research Institute (irri.org)). También se puede emplear SAS (Statistical Analysis System; SAS OnDemand for Academics | SAS) u otros paquetes estadísticos que los generen. Para realizar cálculos manuales con formas cuadráticas o matriciales podría utilizarse la calculadora disponible gratuitamente en <https://matrixcalc.org/es/>. Para analizar datos podrían utilizarse los paquetes estadísticos referenciados previamente u otros como InfoGen y InfoStat (Balzarini *et al.*, 2008; Di Rienzo *et al.*, 2008; Balzarini y Di Rienzo, 2016); éstos últimos se han aplicado para dividir al Error Conjunto en Errores Muestral y Experimental en ensayos con un solo factor, para DCA, DBCA, y DCL (González *et al.*, 2023).

Resultados

En Parcelas Principales, Hileras (H), Columnas (C) y niveles del Factor A son iguales ($h=c=t$); ht ó ct también es igual a t^2 . Para un DBCA, H ó C, pero no ambos, podría considerarse como r, el número de repeticiones elegido en el experimento.

Grados de libertad (GL) sin submuestreo

$$GL \text{ Total} = t^2b - 1$$

$$GL \text{ H} = h-1 = t-1 ; GL \text{ C} = c-1 = t-1 ; GL \text{ A} = t-1 ; GL \text{ Error a} = (t-1)(t-2)$$

$$GL \text{ Parcelas Principales (PP)} = GL \text{ H} + GL \text{ C} + GL \text{ A} + GL \text{ Error a}$$

$$\text{Para verificación: } GL \text{ PP} = t^2 - 1$$

$$GL \text{ B} = b-1 ; GL \text{ AxB} = (t-1)(b-1)$$

$$GL \text{ Error b} = t(t-1)(b-1)$$

$$GL \text{ Subparcelas (SUB)} = GL \text{ Total} - GL \text{ H} - GL \text{ C} - GL \text{ A} - GL \text{ Error a} - GL \text{ B} - GL \text{ AxB}$$

$$\text{También: } GL \text{ SUB} = GL \text{ B} + GL \text{ AxB} + GL \text{ Error b}$$

$$\text{Para verificación: } GL \text{ SUB} = t^2 (b-1)$$

$$GL \text{ PP} + GL \text{ SUB} = GL \text{ Total} = (t^2 - 1) + t^2 (b-1) = t^2b - 1$$

Suma de Cuadrados (SC) sin submuestreo

En el denominador de las siguientes fórmulas, h ó c es nulo; H, C, A, B estarán representados por i, j, k, l , respectivamente. Las formas cuadráticas o matriciales se escribirán como lo hicieron González *et al.* (2023), y González *et al.* (2024 a, b); en éstas se utilizarán sumas o totales realizadas sobre el o los subíndices que no se muestran en su numerador.

$$\text{SC Total} = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}}{t^2 b} \right)^2 = Y'Y - \left(\frac{1}{t^2 b} \right) Y'JY$$

$$\text{SC H} = \left(\frac{1}{tb} \right) \sum_{i=1}^h Y_{i...}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}}{t^2 b} \right)^2 = \left(\frac{1}{tb} \right) Y'_{i...} Y_{i...} - \left(\frac{1}{t^2 b} \right) Y'JY$$

$$\text{SC C} = \left(\frac{1}{tb} \right) \sum_{j=1}^c Y_{.j..}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}}{t^2 b} \right)^2 = \left(\frac{1}{tb} \right) Y'_{.j..} Y_{.j..} - \left(\frac{1}{t^2 b} \right) Y'JY$$

$$\text{SC A} = \left(\frac{1}{tb} \right) \sum_{k=1}^t Y_{..k.}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}}{t^2 b} \right)^2 = \left(\frac{1}{tb} \right) Y'_{..k.} Y_{..k.} - \left(\frac{1}{t^2 b} \right) Y'JY$$

Para calcular SC Error a, primero deberá calcularse SC TRAT1, la cual se obtiene cómo:

$$\text{SC TRAT1} = \left(\frac{1}{b} \right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k.}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}}{t^2 b} \right)^2 = \text{SC H} + \text{SC C} + \text{SC A} + \text{SC Error a}$$

Error a

$$\text{También: SC TRAT1} = \left(\frac{1}{b} \right) Y'_{i.k.} Y_{i.k.} - \left(\frac{1}{t^2 b} \right) Y'JY = \text{SC Parcelas Principales}$$

Por lo tanto: SC Error a = SC TRAT1 – SC H – SC C – SC A

Para verificar que **SC Error a** sea correcta, aplicar la fórmula alternativa:

$$\begin{aligned} \text{SC Error a} &= \left(\frac{1}{b} \right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k.}^2 - \left(\frac{1}{tb} \right) \sum_{i=1}^h Y_{i...}^2 - \left(\frac{1}{tb} \right) \sum_{j=1}^c Y_{.j..}^2 - \left(\frac{1}{tb} \right) \sum_{k=1}^t Y_{..k.}^2 + 2 FC \\ &= \left(\frac{1}{b} \right) Y'_{i.k.} Y_{i.k.} - \left(\frac{1}{tb} \right) Y'_{i...} Y_{i...} - \left(\frac{1}{tb} \right) Y'_{.j..} Y_{.j..} - \left(\frac{1}{tb} \right) Y'_{..k.} Y_{..k.} + \left(\frac{2}{t^2 b} \right) Y'JY \end{aligned}$$

De lo anterior también podrá verificarse que la suma de cuadrados de las Parcelas Principales

(SC PP) es igual a:

$$\text{SC PP} = \left(\frac{1}{b} \right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k.}^2 - FC$$

$$= \left(\frac{1}{b}\right) Y'_{i.k} Y_{i.k} - \left(\frac{1}{t^2 b}\right) Y' J Y$$

$$SC B = \left(\frac{1}{t^2}\right) \sum_{l=1}^b Y_{...l}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}}{t^2 b}\right)^2 = \left(\frac{1}{t^2}\right) Y'_{...l} Y_{...l} - \left(\frac{1}{t^2 b}\right) Y' J Y$$

Para calcular SC AxB, deberá obtenerse primero SC TRAT2 cómo:

$$SC TRAT2 = \left(\frac{1}{t}\right) \sum_{k=1}^t \sum_{l=1}^b Y_{..kl}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}}{t^2 b}\right)^2 = \left(\frac{1}{t}\right) Y'_{..kl} Y_{..kl} - \left(\frac{1}{t^2 b}\right) Y' J Y$$

El denominador de la primera parte de la formula debe ser h ó c, pero no ambos y, como h=c=t, por eso se escribió t, el número de niveles en el Factor A

$$SC TRAT2 = SC A + SC B + SC AxB$$

$$Por\ lo\ tanto: SC AxB = SC TRAT2 - SC A - SC B$$

$$SC Error b = SC Total - SC H - SC C - SC A - SC Error a - SC B - SC AxB$$

Para verificar que el cálculo anterior sea correcto deberá aplicarse la fórmula alternativa:

$$SC Error b = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}^2 - \left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k}^2 - \left(\frac{1}{t}\right) \sum_{k=1}^t \sum_{l=1}^b Y_{..kl}^2 + \left(\frac{1}{tb}\right) \sum_{k=1}^t Y_{..k}^2$$

$$= Y' Y - \left(\frac{1}{b}\right) Y'_{i.k} Y_{i.k} - \left(\frac{1}{t}\right) Y'_{..kl} Y_{..kl} + \left(\frac{1}{tb}\right) Y'_{..k} Y_{..k}$$

También es verificable para Subparcelas (SUB) la fórmula:

$$SC SUB = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}^2 - \left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k}^2$$

$$= Y' Y - \left(\frac{1}{b}\right) Y'_{i.k} Y_{i.k}$$

En el contexto previo deberá verificarse: SC Total = SC PP + SC SUB

Grados de libertad (GL) con submuestreo

Debido a la adición del Factor S ($m=1,2,3,\dots,s$):

$$GL \text{ Total} = t^2bs - 1$$

$$GL \text{ H} = h - 1 = t - 1$$

$$GL \text{ C} = c - 1 = t - 1$$

$$GL \text{ A} = t - 1$$

$$GL \text{ Error a} = (t - 1)(t - 2)$$

$$GL \text{ PP} = GL \text{ H} + GL \text{ C} + GL \text{ A} + GL \text{ Error a}$$

$$\text{También: } GL \text{ PP} = t^2 - 1$$

$$GL \text{ B} = b - 1$$

$$GL \text{ AxB} = (t - 1)(b - 1)$$

$$GL \text{ EC} = GL \text{ Total} - GL \text{ H} - GL \text{ C} - GL \text{ A} - GL \text{ Error a} - GL \text{ B} - GL \text{ AxB}$$

Dónde: GL EC son los grados de libertad del Error Conjunto

$$\text{También : } GL \text{ EC} = t(tbs - t - b + 1)$$

Además, se sabe que (González *et al.*, 2023):

$$GL \text{ EC} = GL \text{ EM} + GL \text{ EE}$$

Pero $GL \text{ EE} = t(t - 1)(b - 1)$, por lo tanto:

$$GL \text{ EM} = GL \text{ EC} - GL \text{ EE} = t^2b(s - 1)$$

Para verificación:

$$GL\ SUB = t^2 (bs - 1)$$

$$GL\ PP + GL\ SUB = (t^2 - 1) + t^2 (bs - 1) = t^2 bs - 1 = GL\ Total$$

Suma de cuadrados (SC) con submuestreo

$$SC\ Total = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs} \right)^2 = Y'Y - \left(\frac{1}{t^2 bs} \right) Y'JY$$

$$SC\ H = \left(\frac{1}{tbs} \right) \sum_{i=1}^h Y_{i\dots}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs} \right)^2 = \left(\frac{1}{tbs} \right) Y'_{i\dots} Y_{i\dots} - \left(\frac{1}{t^2 bs} \right) Y'JY$$

$$SC\ C = \left(\frac{1}{tbs} \right) \sum_{j=1}^c Y_{.j\dots}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs} \right)^2 = \left(\frac{1}{tbs} \right) Y'_{.j\dots} Y_{.j\dots} - \left(\frac{1}{t^2 bs} \right) Y'JY$$

$$SC\ A = \left(\frac{1}{tbs} \right) \sum_{k=1}^t Y_{\dots k\dots}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs} \right)^2 = \left(\frac{1}{tbs} \right) Y'_{\dots k\dots} Y_{\dots k\dots} - \left(\frac{1}{t^2 bs} \right) Y'JY$$

SC PP = SC H + SC C + SC A + SC HxA. En esta, SC HxA = SC Error a

$$Así: SC\ PP = SC\ TRAT\ 1 = \left(\frac{1}{bs} \right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k\dots}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs} \right)^2 = \left(\frac{1}{bs} \right) Y'_{i.k\dots} Y_{i.k\dots} - \left(\frac{1}{t^2 bs} \right) Y'JY$$

$$Por\ lo\ tanto: SC\ HxA = SC\ TRAT\ 1 - SC\ H - SC\ C - SC\ A = \left(\frac{1}{bs} \right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k\dots}^2 -$$

$$\left(\frac{1}{tbs} \right) \sum_{i=1}^h Y_{i\dots}^2 - \left(\frac{1}{tbs} \right) \sum_{j=1}^c Y_{.j\dots}^2 - \left(\frac{1}{tbs} \right) \sum_{k=1}^t Y_{\dots k\dots}^2 + 2 \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs} \right)^2$$

$$También: SC\ HxA = \left(\frac{1}{bs} \right) Y'_{i.k\dots} Y_{i.k\dots} - \left(\frac{1}{tbs} \right) Y'_{i\dots} Y_{i\dots} - \left(\frac{1}{tbs} \right) Y'_{.j\dots} Y_{.j\dots} - \left(\frac{1}{tbs} \right) Y'_{\dots k\dots} Y_{\dots k\dots} +$$

$$\left(\frac{2}{t^2 bs} \right) Y'JY$$

$$SC B = \left(\frac{1}{t^2 s}\right) \sum_{l=1}^b Y_{...l}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs}\right)^2 = \left(\frac{1}{t^2 s}\right) Y'_{...l} Y_{...l} - \left(\frac{1}{t^2 bs}\right) Y' J Y$$

Para calcular SC AxB, primero deberá calcularse SC TRAT2, cómo:

$$SC TRAT2 = \left(\frac{1}{ts}\right) \sum_{k=1}^t \sum_{l=1}^b Y_{..kl}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs}\right)^2 = \left(\frac{1}{ts}\right) Y'_{..kl} Y_{..kl} - \left(\frac{1}{t^2 bs}\right) Y' J Y$$

$$SC AxB = SC TRAT2 - SC A - SC B = \left(\frac{1}{ts}\right) \sum_{k=1}^t \sum_{l=1}^b Y_{..kl}^2 -$$

$$\left(\frac{1}{tbs}\right) \sum_{k=1}^t Y_{..k..}^2 - \left(\frac{1}{t^2 s}\right) \sum_{l=1}^b Y_{...l}^2 + \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs}\right)^2 = \left(\frac{1}{ts}\right) Y'_{..kl} Y_{..kl} -$$

$$\left(\frac{1}{tbs}\right) Y'_{..k..} Y_{..k..} - \left(\frac{1}{t^2 s}\right) Y'_{...l} Y_{...l} + \left(\frac{1}{t^2 bs}\right) Y' J Y$$

Ahora ya se puede obtener, por diferencia, la SC EC, que es la suma de cuadrados del Error

Conjunto. Su valor se estima a partir de:

$$SC EC = SC Total - SC H - SC C - SC A - SC Error a - SC B - SC AxB$$

También es verificable:

$$SC EC = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}^2 - \left(\frac{1}{bs}\right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k..}^2 - \left(\frac{1}{ts}\right) \sum_{k=1}^t \sum_{l=1}^b Y_{..kl}^2 +$$

$$\left(\frac{1}{tbs}\right) \sum_{k=1}^t Y_{..k..}^2 = Y' Y - \left(\frac{1}{bs}\right) Y'_{i.k..} Y_{i.k..} - \left(\frac{1}{ts}\right) Y'_{..kl} Y_{..kl} + \left(\frac{1}{tbs}\right) Y'_{..k..} Y_{..k..}$$

$$SC Error Muestral (EM) = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}^2 - \left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t \sum_{l=1}^b Y_{i.kl}^2$$

$$= Y' Y - \left(\frac{1}{b}\right) Y'_{i.kl} Y_{i.kl}$$

Adicionalmente, la SC del error experimental (SC EE) es igual a:

$$SC EE = \left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t \sum_{l=1}^b Y_{i.kl}^2 - \left(\frac{1}{bs}\right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k..}^2 - \left(\frac{1}{ts}\right) \sum_{k=1}^t \sum_{l=1}^b Y_{..kl}^2 +$$

$$\left(\frac{1}{tbs}\right) \sum_{k=1}^t Y_{..k..}^2$$

$$= \left(\frac{1}{b}\right)Y'_{i.kl}.Y_{i.kl}. - \left(\frac{1}{bs}\right)Y'_{i.k..}Y_{i.k..} - \left(\frac{1}{ts}\right)Y'_{..kl}.Y_{..kl}. + \left(\frac{1}{tbs}\right)Y'_{..k..}Y_{..k..}$$

Discusión

Los arreglos en PD y Parcelas Subdivididas para DCA y DBCA, han sido utilizados más frecuentemente en ensayos conducidos en las ciencias agropecuarias y forestales, entre otras (Gomez y Gomez, 1984; Martínez, 1988, 1994; Montgomery, 2009; González *et al.*, 2019), pero la correspondientes al Cuadro Latino (DCL) no está bien documentado (Sahagún, 1998; Ledolter, 2010; Tirado y Tirado, 2017; González *et al.*, 2022). En este estudio se aplica el DCL a los niveles del Factor A en Parcelas Principales, y los correspondientes al Factor B son asignadas a las Subparcelas de manera completamente al azar, como lo sugirieron Smith (1951), Martínez (1994), y Tirado y Tirado (2017). Esto origina dos tipos de error: uno igual al de un DCL para un experimento de un factor, y otro que es el residual en el modelo (Martínez, 1994; <https://biometrics.ilri.org/Publication/Full%20Text/chapter20.pdf>).

En la literatura publicada se ha encontrado poca evidencia con relación al modelo estadístico de referencia, así como para un Análisis de varianza (Anava) y una comparación de medias (Ledolter, 2010; Tirado y Tirado, 2017; <https://biometrics.ilri.org/Publication/Full%20Text/chapter20.pdf>), especialmente para submuestreo (Gomez y Gomez, 1984; Martínez, 1988; PROC ANOVA: Latin Square Split Plot :: SAS/STAT® 9.22 User's Guide). Martínez (1994) sólo presentó una tabla con las fuentes de variación (FV) y los grados de libertad (GL) que podrían calcularse sin submuestreo. Gomez y Gomez (1984) mostraron dos tablas con las FV y los GL para un DBCA y para un arreglo en PD con submuestreo balanceado, pero para un solo ensayo; adicionalmente, ellos describieron los procedimientos para obtener un Anava en ambas

situaciones, con énfasis en la estimación de los Errores Muestral (EM) y Experimental (EE), que en el presente estudio conforman al Error Conjunto (EC).

En el Sistema para Análisis Estadístico (SAS, Statistical Analysis System) sólo se presentaron los datos proporcionados por Smith (1951; PROC ANOVA: Latin Square Split Plot :: SAS/STAT® 9.22 User's Guide) pero sin submuestreo, así como un código para obtener un Anava con tres tipos de errores; con relación a la presente investigación, CxHxA es igual al Error a, y la suma de HxB y del residual del modelo originan al Error b. Tirado y Tirado (2017) presentaron el modelo estadístico para un PD en DCL sin submuestreo, así como un ejemplo para generar su Anava; ellos también generalizaron este tipo de análisis para un arreglo en Parcelas Subdivididas en DCL sin submuestreo.

En el presente estudio se construyó el modelo estadístico para experimentos en PD en DCL, sin y con submuestreo balanceado, y se generaron las fórmulas para calcular GL y suma de cuadrados (SC) aplicando Mínimos Cuadrados y formas cuadráticas o matriciales, con base en las recomendaciones realizadas para otros estudios por González *et al.* (2023) y González *et al.* (2024 a, b), aunque ellos también destacaron el uso de software.

En esta investigación se consideró que no hay interacción entre hileras (H), columnas ®, y niveles del Factor A, como lo sugirieron Gomez y Gomez (1984), Martínez (1988, 1994), Montgomery (2009), y Tirado y Tirado (2017), entre otros, pero aplicando la guía publicada por Sahagún (1998), se llegará a los mismos resultados con las consideraciones siguientes: H anidado en C, o viceversa; H, C o ambos anidados dentro del Factor A; los factores A y B están cruzados; H, C o ambos anidados en la interacción AxB; el error conjunto anidado en todas las componentes de este modelo. Con submuestreo balanceado, además de lo anterior, S estará anidado en H, C y AxB.

Al aplicar InfoGen o InfoStat será correcto elegir cómo Error a, alguna de las siguientes combinaciones: HxC, CxH, HxA, CxA, o HxCxA, debido a que en un DCL también es cierto que $H=C=A=R$, donde R es el número de repeticiones si se eligiera un DCA o un DBCA. En el presente estudio la interacción HxA fue considerada como Error a, pero al considerar el código para SAS que permitió analizar los datos de Smith (1951) éste es equivalente a la interacción CxHxA. Así, las componentes restantes en el modelo sin submuestreo serán B, AxB, y Error b. Con submuestreo balanceado, el EC debe fraccionarse en Errores Muestral (EM) y Experimental (EE), y los componentes restantes serán los mismos que para el PD en DCL sin submuestreo. En ambos paquetes estadísticos EM será estimado directamente, cuando en especificaciones a los términos del modelo se capture la instrucción $H*A*B>S$, pero $C*A*B>S$ originará los mismos resultados.

Al aplicar InfoStat o InfoGen (Balzarini *et al.*, 2008; Di Rienzo *et al.*, 2008; Balzarini y Di Rienzo, 2016), la diferencia que se origine al considerar o no a S, el número de veces que se aplica submuestreo balanceado, permitirá el cálculo indirecto de SC EM, pero para validar resultados, también podrán emplearse las fórmulas presentadas previamente, así como las que fueron construidas para estimar EC y EE, si el usuario usa Mínimos Cuadrados, formulas cuadráticas o matriciales, o ambas. El paquete estadístico STAR genera un plan de aleatorización para PD en DCL, así como un Anava y una comparación de medias para los Factores A, B y para su interacción, pero en este no está implementada la modalidad de submuestreo. Esta misma situación se ha observado al revisar diversos paquetes estadísticos que frecuentemente se utilizan en las ciencias agropecuarias y forestales, entre otras (González *et al.*, 2019; Pérez *et al.*, 2022). También podría construirse una tabla para la interacción HxAxB;

la diferencia entre SC Total y SC de esta triple interacción producirá SC EM. Así, SC EE sería la diferencia entre SC EC y SC EM.

Para un PD en DCL sin submuestreo, las hipótesis estadísticas relacionadas con H, C, y A serán probadas utilizando al Error α , o la interacción HxA, mientras que las correspondientes al Factor B, y a la interacción AxB serán evaluadas con el error conjunto (Ledolter, 2010; Tirado y Tirado, 2017; PROC ANOVA: Latin Square Split Plot :: SAS/STAT® 9.22 User's Guide). Para el modelo con submuestreo, primero deberá determinarse si EM es significativo: si lo es, éste será utilizado para probar las hipótesis estadísticas del Factor B y de su interacción; si no lo es, ambas fuentes de variación serán evaluadas utilizando el residual del modelo (Ledolter, 2010; Tirado y Tirado, 2017).

Con este mismo enfoque serán realizadas las pruebas de comparación de medias para las componentes dentro de Parcelas Principales y dentro de Subparcelas (Ledolter, 2010; Tirado y Tirado, 2017), pero también podría recurrirse a las recomendaciones proporcionadas por Gomez y Gomez (1984), Little y Hills (2008), y Sahagún (1998), si es que H, C, pero no ambas, no fueran significativas, en cuyo caso ese tipo de comparaciones serán equivalentes a las de un DBCA en PD. Si H y C tampoco son significativas, el usuario tiene la opción de analizar sus datos como un DCA en PD utilizando la misma base de datos que para los casos referenciados previamente (Balzarini *et al.*, 2008; Balzarini y Di Rienzo, 2016; Di Rienzo *et al.*, 2008; Tirado y Tirado, 2017).

Las metodologías presentadas en esta investigación permitirán homologar fórmulas para calcular grados de libertad y suma de cuadrados de manera fácil y confiable, como lo han mostrado para otros estudios González *et al.* (2023), y González *et al.* (2024 a, b); con su simbología y la descrita en Mendenhall (1987), Sahagún (2007), y Montgomery (2009), se

generarán de manera fácil y confiable las formas cuadráticas o matriciales que alimentarán más frecuentemente a las calculadoras de matrices, o a los paquetes estadísticos como SAS, Agrobases, SPSS, StatGraphics, STAR, PB Tools, entre otros, por lo que los procedimientos para construir un modelo estadístico, para obtener una Anava y para comparar medias de tratamientos en experimentos factoriales, para el tipo de arreglo de unidades experimentales aquí considerado, permitirán verificar directa o indirectamente, los cálculos pertinentes a ambas metodologías de manera más rápida y segura.

Conclusiones

En los dos modelos estadísticos que se construyeron en esta investigación se asumió que no hay interacción entre Hileras, Columnas y niveles del Factor A, asignables a Parcelas Principales en un diseño en Cuadro Latino; adicionalmente se consideró que hay cruzamiento entre los factores A y B. Ambos modelos también se pueden generar si se considera lo siguiente: H anidado en C, o viceversa; H, C o ambos anidados dentro del Factor A; los factores A y B están cruzados; H, C o ambos están anidados en la interacción AxB; el error conjunto se anida en todas las componentes de este modelo. Con submuestreo balanceado, además de lo anterior, S estará anidado en H, C y AxB. Para verificar grados de libertad y sumas de cuadrados en ambos modelos pueden aplicarse las fórmulas alternativas que fueron construidas para los Errores a, Conjunto, Muestral y Experimental, para cada una o para ambas metodologías. Otra opción sería aplicar directamente las fórmulas generadas para Parcelas Principales y Subparcelas, y estimar indirectamente los cuatro errores de referencia.

Literatura citada

Balzarini, M. G. y Di Rienzo, J. A. 2016. InfoGen. FCA. Universidad Nacional de Córdoba, Argentina (<http://www.info-Gen.com.ar>).

Balzarini, M. G.; González, L.; Tablada, M.; Casanoves, F.; Di Rienzo, J. A. y Robledo, C. W. 2008. Manual del usuario de InfoStat. Ed. Brujas, Córdoba, Argentina. 348 pp.

Di Rienzo, J. A.; Casanoves, F.; Balzarini, M. G.; González, L.; Tablada, M.; Robledo, C. W. 2008. InfoStat, versión 2008. Grupo InfoStat, FCA. Universidad Nacional de Córdoba, Argentina. (<https://www.infostat.com.ar>).

Gomez, K. A. y Gomez, A. A. 1984. Statistical procedures for agricultural research. 2nd Ed. John Wiley & Sons, Inc. Printed in Singapore. 680 pp.

González, H. A.; Pérez, L. D. J.; Rubí, A. M.; Gutiérrez, R. F.; Franco, M. J. R. P. y Padilla, L. A. 2019. InfoStat, InfoGen y SAS para contrastes mutuamente ortogonales en experimentos en bloques completos al azar en parcelas subdivididas. *Revista Mexicana de Ciencias Agrícolas*. 10(6):1417-1431.

González, H. A.; Pérez, L. D. J.; Balbuena, M. A.; Franco, M. J. R.; Gutiérrez, R. F. y Rodríguez, G. J. A. 2023. Submuestreo balanceado en experimentos monofactoriales usando InfoStat y InfoGen: validación con SAS. *Revista Mexicana de Ciencias Agrícolas*. 14(2):235-249 (DOI:<https://doi.org/10.29312/remexca.V14i2.3418>).

González, H., A.; Pérez, L., D.J.; Hernández, A., J.; Franco, M., J.R.P.; Rubí, A., M.; Balbuena, M., A. 2024 a. Tratamientos anidados dentro de un arreglo en grupos de bloques completos balanceados. *Revista Mexicana de Ciencias Agrícolas*.15(2), e 3634 (DOI: <https://doi.org/10.29312/remexca.v15i2.3634>).

González, H., A.; Pérez, L., D.J.; Hernández, A., J.; Franco, M., J.R.P.; Balbuena, M., A.; Rubí, A., M. 2024 b. Serie de experimentos para tratamientos anidados en grupos en arreglo de bloques completos balanceados. *Revista Mexicana de Ciencias Agrícolas*. 15(7) (En edición).

Ledolter, J. 2010. Split-plot design: 69discusión and examples. *International Journal of Quality Engineering and Technology*. 1(4): 441-457 (DOI: <https://doi.org/10.1504/IJQET.2010.035588>).

Little, T. M. y Hills, F. J. 2008. *Métodos Estadísticos para la Investigación en la Agricultura*. Ed. Trillas, SA de CV. México, DF. 270 p. ISBN:978-968-24-3629-1.

Martínez, G. A. 1988. *Diseños experimentales. Métodos y elementos de teoría*. Editorial. Trillas, 1ª Ed. México, DF. 756 pp.

Martínez, G., A. 1994. *Experimentación Agrícola: Métodos Estadísticos*. Primera Edición. Universidad Autónoma Chapingo, México, D.F., 357 p (ISBN: 9789 6888 42 867).

Mendenhall, W. 1987. *Introducción a la Probabilidad y la Estadística*. Grupo Editorial Iberoamérica. 1ª Ed. México, DF. 626 pp.

Montgomery, D. C. 2009. *Design and Analysis of Experiments*. 7th Ed. John Wiley & Sons, Inc. U.S.A. 656 pp.

Pérez, L., D.; Jasso, B., G.; Saavedra, G., C.; Franco, M., J.R.P.; Ramírez, D., J. F.; González, H., A. 2022. Uso de artificios en Opstat para analizar series de experimentos en dialélico parcial. *Revista Mexicana de Ciencias Agrícolas*. 13(2) 273-287 (DOI: <https://doi.org/10.29312/remexca.v13i2.3130>).

Sahagún, C. J. 1998. Construcción y análisis de los modelos fijos, aleatorios y mixtos. Departamento de Fitotecnia. Programa Nacional de Investigación en Olericultura. Universidad Autónoma Chapingo (UACH). Boletín Técnico Núm. 2, 64 p.

Sahagún, C. J. 2007. Estadística Descriptiva y Probabilidad: Una Perspectiva Biológica. 2ª Ed. Universidad Autónoma Chapingo (UACH). México, D.F. 282 pp. ISBN: 978 -968-02- 0357-4.

Smith, W., G. 1951. Dissertation notes on Canadian Sugar Factories Ltd, Taber, Alberta, Canadá.

Tirado, E., G., y Tirado, G., D. N. 2017. Tratado de Estadística Experimental. Primera Edición. Editorial Centro de Estudios e Investigaciones para el Desarrollo Docente (CENID, AC), Guadalajara, Jalisco, México (ISBN: 978-607-8435-43-2).

4.3 Capitulo de libro

Un ejemplo para parcelas divididas en cuadro latino, sin y con submuestreo balanceado

Se publicará en la obra **“Tópicos Selectos para Diseñar y Aplicar Tecnología Agropecuaria”**.

Libro que forma parte de las actividades de vinculación entre el Cuerpo Académico “CA 127 UAEM Producción de Cultivos Básicos y Hortícolas”, de la Facultad de Ciencias Agrícolas, y otras CA con registro UAEM-SEP.

Un ejemplo para Parcelas Divididas en Cuadro Latino, sin y con submuestreo balanceado

José Antonio Rodríguez González¹, Delfina de Jesús Pérez López², Artemio Balbuena Melgarejo², Jesús Hernández Ávila², J. Ramón Pascual Franco Martínez², Laura Stephanie Flores Carrera¹, Claudia Saavedra Guevara¹, Gerardo Jasso Bobadilla¹, y Andrés González Huerta^{2δ}

¹Programa de Maestría y Doctorado en Ciencias Agropecuarias y Recursos Naturales, Universidad Autónoma del Estado de México (UAEMéx), Facultad de Ciencias Agrícolas (FCAgr), Campus Universitario “El Cerrillo”. El Cerrillo Piedras Blancas, Municipio de Toluca de Lerdo, Estado de México, México (CUC-CPB-MTLM). Tel. + 52 (722) 29 655 52. Ext. 117 (pcarn@uaemex.mx). ²Centro de Investigación y Estudios Avanzados en Fitomejoramiento (C.I.E.A.F.), FCAgr - UAEMéx. CUC-CPB-MTLM. Apartado Postal 435, Tel: +52 722 29 655 18, Ext. 148 (jrodriguezg008@alumno.uaemex.mx; djperezl@uaemex.mx; jhernandez@uaemex.mx; jrfrancom@uaemex.mx; abalbuenam@uaemex.mx. □ Autor para correspondencia: agonzalezh@uaemex.mx

Resumen

Debido a la escasa información que se ha encontrado en la literatura reportada para un arreglo en parcelas divididas en diseño en cuadro latino, sin y con submuestreo balanceado, en esta contribución se presenta un ejemplo para generar un análisis de varianza y una comparación de medias para ambos factores y para la combinación entre los niveles de cada factor aplicando la prueba de la diferencia mínima significativa de Fisher obtenidas con el paquete estadístico InfoStat; sólo se presentan los cálculos manuales para generar grados de libertad y suma de

cuadrados cuando se utilizan las fórmulas derivadas por Rodríguez et al. (2025) a datos provenientes de un solo ensayo, pero éstos se pueden extender a una serie de experimentos si fueran empleadas las fórmulas presentadas por Pérez et al. (2025). El modelo estadístico correspondiente al caso de sin submuestreo estadístico difiere sólo en una componente con relación a un arreglo en franjas o bloques divididos en cuadro latino.

Palabras clave: Mínimos Cuadrados, formas cuadráticas o matriciales, experimentos bifactoriales.

Introducción

Los arreglos de unidades experimentales en Parcelas Divididas (PD) o Subdivididas (PSD) han sido utilizados ampliamente para estudiar efectos o varianzas en ensayos bi o trifactoriales (Balzarini *et al.*, 2008; Di Rienzo *et al.*, 2008; Balzarini y Di Rienzo, 2016; Tirado y Tirado, 2017), o en series de experimentos cuando éstos se extienden a diversos ambientes, conformados por años, localidades o combinaciones de ambos (González *et al.*, 2019; Pérez *et al.*, 2022). Los planes de aleatorización más utilizados corresponden a los diseños experimentales Completamente al Azar (DCA) y Bloques Completos al Azar (DBCA) (Martínez, 1988; Gomez y Gomez, 1984; Little y Hills, 2008; Montgomery, 2009). Ledolter (2010) realizó una discusión acerca de los arreglos de PD para ensayos completos y para experimentos factoriales fraccionados, con énfasis en DCA y DBCA, pero sin considerar submuestreo balanceado.

El DCL, en PD o PSD, no ha sido muy utilizado, pero podría ser de gran utilidad para evaluar las diferencias que originan años, localidades o ambos con distanciamiento entre plantas, fórmulas de fertilización, abonos orgánicos, insecticidas, fungicidas, herbicidas, métodos de labranza, láminas de riego, y fechas de corte en forrajes, entre otros, cuando éstos se asignen a

parcela principal, y en subparcelas podrían considerarse especies vegetales o variedades contrastantes de alguna de éstas (González *et al.*, 2019; Pérez *et al.*, 2022). Los terrenos de agricultores que son muy heterogéneos también presentan este tipo de variabilidad aleatoria que es indeseable para establecer un experimento adecuado; ésta podría controlarse más eficientemente utilizando un PD en DCL u otro diseño experimental (González *et al.*, 2019; Pérez *et al.*, 2022).

También se han encontrado pocos reportes sobre los modelos estadísticos, el análisis de varianza y la comparación de medias (Tirado y Tirado, 2017; González *et al.*, 2019), particularmente cuando se aplica submuestreo balanceado y algún paquete estadístico (Gomez y Gomez, 1984; Martínez, 1994; Montgomery, 2009; PROC ANOVA: Latin Square Split Plot :: SAS/STAT(R) 9.22 User's Guide). Así, el objetivo principal de esta investigación fue construir sus modelos sin y con submuestreo balanceado, y presentar las fórmulas para calcular grados de libertad y suma de cuadrados con base en dos metodologías, como un prerrequisito para la aplicación de algún paquete estadístico.

Materiales y métodos

Simbología

En este estudio la variable cuantitativa de interés será identificada con Y. Adicionalmente, será empleada la simbología descrita por Mendenhall (1987), Sahagún (2007) y Montgomery (2009). También será aplicada la terminología descrita por Pérez *et al.* (2022) y por González *et al.* (2023, 2004 a, b). Los factores de clasificación serán Hileras, Columnas, A, B, cuyos niveles serán, respectivamente: $i=1,2,3,\dots,h$; $j=1,2,3,\dots,c$; $k=1,2,3,\dots,t$; $l=1,2,3,\dots,b$. Con submuestreo balanceado, además de lo anterior, será incluido el Factor S ($m=1,2,3,\dots,s$).

Modelos estadísticos balanceados

Los dos modelos que se presentan a continuación se construyeron aplicando la guía publicada por Sahagún (1998), pero ambos también se pueden construir con la suposición de ausencia de interacción entre Hileras, Columnas y niveles del Factor A, como lo sugirieron Gomez y Gomez (1984), Martínez (1994), y Montgomery (2009). Ambos fueron reportados en Rodríguez *et al.* (2025).

Sin submuestreo:

$$Y_{ijkl} = \mu + H_i + C_j + A_k + (HA)_{ik} + B_l + (AB)_{kl} + \varepsilon_{ijkl}$$

Con submuestreo:

$$Y_{ijklm} = \mu + H_i + C_j + A_k + (HA)_{ik} + B_l + (AB)_{kl} + \delta_{m(ikl)} + \varepsilon_{ijklm}$$

Dónde: Y es la variable de interés; μ es la media aritmética general; H_i , C_j , A_k , B_l , son los efectos causados por Hileras, Columnas, y Factores A y B, respectivamente; $(HA)_{ik}$ es el Error a; $(AB)_{kl}$ es la interacción entre ambos factores; $\delta_{m(ikl)}$ es el error muestral; ε_{ijkl} y ε_{ijklm} son el error experimental, sin y con submuestreo.

Diseño experimental

Los niveles del Factor A serán asignados a Parcelas Principales con base en un diseño en Cuadro Latino, como lo sugirieron Smith (1951), Martínez (1994) y Tirado y Tirado (2017), y los niveles del Factor B se aleatorizan en Subparcelas de manera completamente al azar; en un modelo no se considerará submuestreo balanceado y en otro sí.

Software recomendable

Para obtener un plan de aleatorización para este tipo de experimento se puede utilizar el paquete estadístico STAR, del Instituto Internacional de Investigación en Arroz (IRRI), con sede en los Baños, Filipinas (Digital Tools | International Rice Research Institute (irri.org)). También se puede emplear SAS (Statistical Analysis System; SAS OnDemand for Academics | SAS) u otros paquetes estadísticos que los generen. Para realizar cálculos manuales con formas cuadráticas o matriciales podría utilizarse la calculadora disponible gratuitamente en <https://matrixcalc.org/es/>. Para analizar datos podrían utilizarse los paquetes estadísticos referenciados previamente u otros como InfoGen y InfoStat (Balzarini *et al.*, 2008; Di Rienzo *et al.*, 2008; Balzarini y Di Rienzo, 2016); éstos últimos se han aplicado para dividir al Error Conjunto en Errores Muestral y Experimental en ensayos con un solo factor, para DCA, DBCA, y DCL (González *et al.*, 2023).

Resultados

En Parcelas Principales, Hileras (H), Columnas (C) y niveles del Factor A son iguales ($h=c=t$); ht ó ct también es igual a t^2 . Para un DBCA, H ó C, pero no ambos, podría considerarse como r , el número de repeticiones elegido en el experimento.

Grados de libertad (GL) sin submuestreo

$$GL \text{ Total} = t^2b - 1 = 4(4)(3) - 1 = 47$$

$$GL \text{ H} = h-1 = t-1; GL \text{ C} = c-1 = t-1 = GL \text{ A} = t-1 = 4 - 1 = 3; GL \text{ Error a} = (t-1)(t-2) = 3(2) = 6$$

$$GL \text{ Parcelas Principales (PP)} = GL \text{ H} + GL \text{ C} + GL \text{ A} + GL \text{ Error a} = 3 + 3 + 3 + 6 = 15$$

$$\text{Para verificación: } GL \text{ PP} = t^2 - 1 = 4(4) - 1 = 15$$

$$GL \text{ B} = b-1 = 3 - 1 = 2; GL \text{ AxB} = (t-1)(b-1) = 3(2) = 6$$

$$GL \text{ Error } b = t(t-1)(b-1) = 4(3)(2) = 24$$

$$GL \text{ Subparcelas (SUB)} = GL \text{ Total} - GL \text{ H} - GL \text{ C} - GL \text{ A} - GL \text{ Error a} - GL \text{ B} - GL \text{ AxB} = 47 - 3 - 3 - 3 - 6 - 2 - 6 = 24$$

$$\text{También: } GL \text{ SUB} = GL \text{ B} + GL \text{ AxB} + GL \text{ Error } b = 2 + 6 + 24 = 32$$

$$\text{Para verificación: } GL \text{ SUB} = t^2 (b-1) = 4(4)(2) = 32$$

$$GL \text{ PP} + GL \text{ SUB} = GL \text{ Total} = (t^2 - 1) + t^2 (b-1) = t^2 b - 1 = 15 + 32 = 47$$

Suma de Cuadrados (SC) sin submuestreo

En el denominador de las siguientes fórmulas, h ó c es nulo; H, C, A, B estarán representados por i, j, k, l , respectivamente. Las formas cuadráticas o matriciales se escribirán como lo hicieron González *et al.* (2023), y González *et al.* (2024 a, b); en éstas se utilizarán sumas o totales realizadas sobre el o los subíndices que no se muestran en su numerador.

$$SC \text{ Total} = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}}{t^2 b} \right)^2 = (7.6^2 + 2.1^2 + 6.0^2 + \dots, 8.0^2) - \left(\frac{1}{4(4)(3)} \right) (7.6 + 2.1 + 6.0 + \dots, 8.0)^2 = 2276.5 - 1981.47 = 295.03$$

$$SC \text{ H} = \left(\frac{1}{tb} \right) \sum_{i=1}^h Y_{i...}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}}{t^2 b} \right)^2 = \left(\frac{1}{4(3)} \right) (75.3^2 + 77.1^2 + 78.5^2 + 77.5^2) - \left(\frac{1}{4(4)(3)} \right) (7.6 + 2.1 + 6.0 + \dots, 8.0)^2 = 1981.9166 - 1981.47 = 0.4466$$

$$SC \text{ C} = \left(\frac{1}{tb} \right) \sum_{j=1}^c Y_{.j..}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}}{t^2 b} \right)^2 = \left(\frac{1}{4(3)} \right) (77.1^2 + 76.6^2 + 77.1^2 + 77.6^2) - \left(\frac{1}{4(4)(3)} \right) (7.6 + 2.1 + 6.0 + \dots, 8.0)^2 = 1981.5116 - 1981.47 = 0.04166$$

$$SC A = \left(\frac{1}{tb}\right) \sum_{k=1}^t Y_{..k}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}}{t^2 b}\right)^2 = \left(\frac{1}{4(3)}\right) (66.0^2 + 72.1^2 + 78.0^2 + 92.3^2) -$$

$$\left(\frac{1}{4(4)(3)}\right) (7.6 + 2.1 + 6.0 + \dots, 8.0)^2 = 2013.1416 - 1981.47 = 31.6716$$

Para calcular SC Error a, primero deberá calcularse SC TRAT1 utilizando la información de la Tabla 1.

Tabla 1. Valores para calcular SC para Parcelas Principales

Factor H (i niveles)	Factor A (k niveles)				
	1	2	3	4	Suma
1	15.7	17.9	19.2	22.5	75.3
2	16.6	17.7	19.9	22.9	77.1
3	16.8	18.3	20.1	23.3	78.5
4	16.9	18.2	18.8	23.6	77.5
Suma	66.0	72.1	78.0	92.3	308.4

$$SC TRAT1 = \left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}}{t^2 b}\right)^2 = SC H + SC C + SC A + SC$$

Error a

$$= \left(\frac{1}{3}\right) (15.7^2 + 17.9^2 + 19.2^2 + \dots, + 23.6^2) - \left(\frac{1}{4(4)(3)}\right) (7.6 + 2.1 + 6.0 + \dots, 8.0)^2 = 2014.1133 -$$

1981.47

$$= 32.6433$$

$$SC Error a = SC TRAT1 - SC H - SC C - SC A = 32.6433 - 0.4466 - 0.04166 - 31.6716 =$$

0.48344

Para verificar que **SC Error a** sea correcta, aplicar la fórmula alternativa:

$$\text{SC Error a} = \left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k}^2 - \left(\frac{1}{tb}\right) \sum_{i=1}^h Y_{i..}^2 - \left(\frac{1}{tb}\right) \sum_{j=1}^c Y_{.j.}^2 - \left(\frac{1}{tb}\right) \sum_{k=1}^t Y_{..k}^2 + 2 FC$$

$$= 2014.1133 - 1981.9611 - 1981.5116 - 2013.1416 + 2(1981.47) = 0.4835$$

De lo anterior también podrá verificarse que la suma de cuadrados de las Parcelas Principales (SC PP) es igual a:

$$\text{SC PP} = \left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k}^2 - FC = 2014.1133 - 1981.47 = 32.6433$$

$$\text{SC B} = \left(\frac{1}{t^2}\right) \sum_{l=1}^b Y_{...l}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^c \sum_{l=1}^b Y_{ijkl}}{t^2 b}\right)^2 = \left(\frac{1}{4(4)}\right) (52.0^2 + 116.2^2 + 140.2^2) -$$

$$\left(\frac{1}{4(4)(3)}\right) (7.6 + 2.1 + 6.0 + \dots, 8.0)^2 = 2241.405 - 1981.47 = 259.935$$

Para calcular SC AxB, deberá obtenerse primero SC TRAT2 utilizando la información de la Tabla 2.

Tabla 2. Datos para calcular SC TRAT2

Factor A (k)	Factor B (l)			
	1	2	3	Suma
1	9.4	25.6	31.0	66.0
2	11.6	27.9	32.6	72.1
3	12.7	29.3	36.0	78.0
4	18.3	33.4	40.6	92.3
Suma	52.0	116.2	140.2	308.4

$$\text{SC TRAT2} = \left(\frac{1}{t}\right) \sum_{k=1}^t \sum_{l=1}^b Y_{..kl}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^c \sum_{l=1}^b Y_{ijkl}}{t^2 b}\right)^2 = \left(\frac{1}{4}\right) (9.4^2 + 25.6^2 + 31.0^2$$

$$+ \dots, 40.6^2) -$$

$$\left(\frac{1}{4(4)(3)}\right) (7.6 + 2.1 + 6.0 + \dots, 8.0)^2 = 2273.76 - 1981.47 = 292.29$$

$$SC_{A \times B} = SC_{TRAT2} - SC_A - SC_B = 292.29 - 31.6716 - 259.935 = 0.6834$$

$$SC_{Error\ b} = SC_{Total} - SC_H - SC_C - SC_A - SC_{Error\ a} - SC_B - SC_{A \times B}$$

$$= 295.03 - 0.4466 - 0.04166 - 31.6716 - 0.48344 - 259.935 - 0.6834 = 1.7683$$

Para verificar que el cálculo anterior sea correcto deberá aplicarse la fórmula alternativa:

$$SC_{Error\ b} = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}^2 - \left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k.}^2 - \left(\frac{1}{t}\right) \sum_{k=1}^t \sum_{l=1}^b Y_{..kl}^2 +$$

$$\left(\frac{1}{tb}\right) \sum_{k=1}^t Y_{..k.}^2$$

$$= 2276.5 - 2014.1133 - 2273.76 + 2013.1416 = 1.7683$$

También es verificable para Subparcela (SUB) aplicar la fórmula:

$$SC_{SUB} = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b Y_{ijkl}^2 - \left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k.}^2$$

$$= 2276.5 - 2014.1133 = 262.3867$$

En el contexto previo deberá verificarse: $SC_{Total} = SC_{PP} + SC_{SUB}$

Grados de libertad (GL) con submuestreo

Debido a la adición del Factor S ($m=1,2,3,\dots,s$):

$$GL_{Total} = t^2bs - 1 = 4^2(3)(3) - 1 = 143$$

$$GL_H = h - 1 = t - 1 = 4 - 1 = 3$$

$$GL_C = c - 1 = t - 1 = 4 - 1 = 3$$

$$GL_A = t - 1 = 4 - 1 = 3$$

$$GL \text{ Error a} = (t - 1)(t - 2) = 3(2) = 6$$

$$GL \text{ PP} = GL \text{ H} + GL \text{ C} + GL \text{ A} + GL \text{ Error a} = 3 + 3 + 3 + 6 = 15$$

$$\text{También: } GL \text{ PP} = t^2 - 1 = 4^2 - 1 = 15$$

$$GL \text{ B} = b - 1 = 3 - 1 = 2$$

$$GL \text{ AxB} = (t - 1)(b - 1) = 3(2) = 6$$

$$GL \text{ EC} = GL \text{ Total} - GL \text{ H} - GL \text{ C} - GL \text{ A} - GL \text{ Error a} - GL \text{ B} - GL \text{ AxB} = 143 - 3 - 3 - 3 - 6 - 2 - 6 = 120$$

Dónde: GL EC son los grados de libertad del Error Conjunto

$$\text{También: } GL \text{ EC} = t(tbs - t - b + 1) = 4(36 - 4 - 3 + 1) = 120$$

Además, se sabe que (González *et al.*, 2023):

$$GL \text{ EC} = GL \text{ EM} + GL \text{ EE}$$

Pero $GL \text{ EE} = t(t - 1)(b - 1)$, por lo tanto:

$$GL \text{ EM} = GL \text{ EC} - GL \text{ EE} = t^2b(s - 1) = 4^2(3)(2) = 96$$

Para verificación:

$$GL \text{ SUB} = t^2(bs - 1) = 4^2(8) = 128$$

$$GL \text{ PP} + GL \text{ SUB} = (t^2 - 1) + t^2(bs - 1) = t^2bs - 1 = GL \text{ Total} = 15 + 128 = 143$$

Suma de cuadrados (SC) con submuestreo

$$SC \text{ Total} = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2bs} \right)^2$$

$$= (7.6^2 + 7.0^2 + 8.0^2 + \dots, 8.3^2) - \left(\frac{1}{4(4)(3)(3)} \right) (7.6 + 7.0 + 8.0 + \dots, 8.3)^2 = 6733.5 - 5849.7 = 883.799$$

$$SC H = \left(\frac{1}{tbs} \right) \sum_{i=1}^h Y_{i\dots}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs} \right)^2 =$$

$$\left(\frac{1}{4(3)(3)} \right) (223.8^2 + 230.6^2 + 234.0^2 + 229.4^2) - \left(\frac{1}{4(4)(3)(3)} \right) (7.6 + 7.0 + 8.0 + \dots, 8.3)^2$$

$$= 5851.198889 - 5849.70 = 1.49888$$

$$SC C = \left(\frac{1}{tbs} \right) \sum_{j=1}^c Y_{j\dots}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs} \right)^2 =$$

$$\left(\frac{1}{4(3)(3)} \right) (229.6^2 + 226.9^2 + 229.1^2 + 232.1^2) - \left(\frac{1}{4(4)(3)(3)} \right) (7.6 + 7.0 + 8.0 + \dots, 8.3)^2$$

$$= 5850.095 - 5849.70 = 0.3947$$

$$SC A = \left(\frac{1}{tbs} \right) \sum_{k=1}^t Y_{..k\dots}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs} \right)^2 =$$

$$\left(\frac{1}{4(3)(3)} \right) (197.3^2 + 213.8^2 + 233.0^2 + 273.7^2) - \left(\frac{1}{4(4)(3)(3)} \right) (7.6 + 7.0 + 8.0 + \dots, 8.3)^2$$

$$= 5939.95611 - 5849.70 = 90.256$$

Para calcular suma de cuadrados de Parcelas Principales (SC PP) deberá construirse la siguiente tabla:

Tabla X. Datos para calcular suma de cuadrados para Parcelas Principales (SC PP)

Factor H (i niveles)	Factor A (k niveles)				
	1	2	3	4	Suma
1	47.0	52.5	57.3	67.0	223.8
2	49.6	53.0	59.7	68.3	230.6
3	50.0	54.8	60.2	69.0	234.0
4	50.7	53.5	55.8	69.4	229.4
Suma	197.3	213.8	233.0	273.7	917.8

SC PP = SC H + SC C + SC A + SC HxA. En esta, SC HxA = SC Error a

$$\begin{aligned} \text{Así: SC PP} &= \text{SC TRAT 1} = \left(\frac{1}{bs}\right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k..}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs}\right)^2 \\ &= \left(\frac{1}{3(3)}\right) (47.0^2 + 52.5^2 + 57.3^2 + \dots + 69.4^2) - \left(\frac{1}{4(4)(3)(3)}\right) (7.6 + 7.0 + 8.0 + \dots + 8.3)^2 \\ &= 5942.9444 - 5849.70 = 93.24 \end{aligned}$$

$$\text{SC HxA} = \text{SC TRAT 1} - \text{SC H} - \text{SC C} - \text{SC A} = 93.24 - 1.49888 - 0.3947 - 90.2561 = 1.09$$

También puede verificarse que:

$$\begin{aligned} \text{SC HxA} &= \text{SC Error a} = \left(\frac{1}{bs}\right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k..}^2 - \left(\frac{1}{tbs}\right) \sum_{i=1}^h Y_{i....}^2 - \\ &\left(\frac{1}{tbs}\right) \sum_{j=1}^c Y_{.j...}^2 - \left(\frac{1}{tbs}\right) \sum_{k=1}^t Y_{..k..}^2 + 2\left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs}\right)^2 \\ &= 5942.9444 - 5851.198889 - 5850.095 - 5939.95611 + 2(5849.70) = 17642.3444 - 17641.25 \\ &= 1.094 \end{aligned}$$

$$\begin{aligned} \text{SC B} &= \left(\frac{1}{t^2s}\right) \sum_{l=1}^b Y_{...l}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs}\right)^2 \\ &= \left(\frac{1}{4(4)(3)}\right) (154.2^2 + 347.3^2 + 416.3^2) - \left(\frac{1}{4(4)(3)(3)}\right) (7.6 + 7.0 + 8.0 + \dots + 8.3)^2 \end{aligned}$$

$$= 6618.7629 - 5849.70 = 769.06$$

Para calcular SC AxB, primero deberá elaborarse la tabla siguiente:

Tabla Y. Datos para generar SC A x B

Factor A (k)	Factor B (l)			
	1	2	3	Suma
1	28.4	76.4	92.5	197.3
2	33.7	83.3	96.8	213.8
3	37.9	87.8	107.3	233.0
4	54.2	99.8	119.7	273.7
Suma	154.2	347.3	416.3	917.8

La interacción AxB se define en los siguientes términos:

SC TRAT2 = SC A + SC B + SC A x B, dónde:

$$SC\ TRAT2 = \left(\frac{1}{ts}\right) \sum_{k=1}^t \sum_{l=1}^b Y_{..kl}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2bs}\right)^2$$

$$= \left(\frac{1}{4(3)}\right) (28.4^2 + 76.4^2 + 92.5^2 + \dots, 119.7^2) - \left(\frac{1}{4(4)(3)(3)}\right) (7.6 + 7.0 + 8.0 + \dots, 8.3)^2$$

$$= 6710.741667 - 5849.70 = 861.041$$

Por cálculo indirecto : SC AxB = SC TRAT2 – SC A – SC B = 861.041 – 90.2561 – 769.06 =

1.72

Por cálculo directo : SC AxB = SC TRAT2 – SC A – SC B = $\left(\frac{1}{ts}\right) \sum_{k=1}^t \sum_{l=1}^b Y_{..kl}^2 -$

$$\left(\frac{1}{tbs}\right) \sum_{k=1}^t Y_{..k..}^2 - \left(\frac{1}{t^2s}\right) \sum_{l=1}^b Y_{...l.}^2 + \left(\frac{\sum_{i=1}^h \sum_{j=1}^t \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2bs}\right)^2$$

$$SC_{A \times B} = 6710.741667 - 5939.95611 - 6618.7629 + 5849.70 = 1.7226$$

Para calcular la suma de cuadrados del Error Conjunto (SC EC) se utiliza la siguiente expresión :

$$SC_{EC} = SC_{Total} - SC_H - SC_C - SC_A - SC_{Error\ a} - SC_B - SC_{A \times B}$$

$$= 883.799 - 1.49888 - 0.3947 - 90.2561 - 1.09 - 769.06 - 1.72 = 19.779$$

También es verificable :

$$SC_{EC} = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}^2 - \left(\frac{1}{bs}\right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k..}^2 - \left(\frac{1}{ts}\right) \sum_{k=1}^t \sum_{l=1}^b Y_{..kl.}^2 +$$

$$\left(\frac{1}{tbs}\right) \sum_{k=1}^t Y_{..k..}^2 = 6733.5 - 5942.9444 - 6710.741667 + 5939.95611 = 19.77$$

De manera similar a casos anteriores, para calcular la suma de cuadrados del error muestral (SC EM) y la del error experimental (SC EE) deberá construirse la tabla siguiente :

(i)	Combinación de niveles en A y B (kl)												
	11	12	13	21	22	23	31	32	33	41	42	43	Suma
1	6.3	18.1	22.6	7.9	20.4	24.2	9.3	21.9	26.1	12.5	24.4	30.1	223.8
2	6.9	18.9	23.8	7.8	21.2	24.0	9.7	22.8	27.2	14.3	25.5	28.5	230.6
3	7.9	20.0	22.1	10.1	19.7	25.0	10.1	22.2	27.9	13.5	24.2	31.3	234.0
4	7.3	19.4	24.0	7.9	22.0	23.6	8.8	20.9	26.1	13.9	25.7	29.8	229.4
Suma	28.4	76.4	92.5	33.7	83.3	96.8	37.9	87.8	107.3	54.2	99.8	119.7	917.8

En el contexto previo, $SC_{EM} = SC_{Total} - SC_{TRAT3}$

$$Pero\ SC_{TRAT3} = \left(\frac{1}{s}\right) \sum_{i=1}^h \sum_{k=1}^t \sum_{l=1}^b Y_{i.kl.}^2 - \left(\frac{\sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}}{t^2 bs}\right)^2$$

$$= \left(\frac{1}{3}\right) (6.3^2 + 18.1^2 + 22.6^2 + \dots, 31.3^2) - \left(\frac{1}{4(4)(3)(3)}\right) (7.6 + 7.0 + 8.0 + \dots, 8.3)^2$$

$$= 6719.4933 - 5849.70 = 869.793$$

Así: SC EM = 883.799 – 869.793 = 14.006

Adicionalmente: SC EE = SC EC – SC EM = 19.77 – 14.006 = 5.76

Otra ruta para calcular la componente anterior sería:

$$SC EE = SC TRAT3 - SC H - SC C - SC A - SC Error a - SC B - SC Ax B$$

$$= 869.793 - 1.49888 - 0.3947 - 90.2561 - 1.09 - 769.06 - 1.72 = 5.773$$

$$SC Error Muestral (EM) = \sum_{i=1}^h \sum_{j=1}^c \sum_{k=1}^t \sum_{l=1}^b \sum_{m=1}^s Y_{ijklm}^2 - \left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t \sum_{l=1}^b Y_{i.kl}^2.$$

$$= 6733.5 - 6719.4933 = 14.0067$$

Adicionalmente, la SC del error experimental (SC EE) es igual a:

$$SC EE = \left(\frac{1}{b}\right) \sum_{i=1}^h \sum_{k=1}^t \sum_{l=1}^b Y_{i.kl}^2 - \left(\frac{1}{bs}\right) \sum_{i=1}^h \sum_{k=1}^t Y_{i.k..}^2 - \left(\frac{1}{ts}\right) \sum_{k=1}^t \sum_{l=1}^b Y_{..kl}^2 +$$

$$\left(\frac{1}{tbs}\right) \sum_{k=1}^t Y_{..k..}^2.$$

$$= 6719.4933 - 5942.9444 - 6710.741667 + 5939.95611 = 5.7633$$

Aplicación de InfoStat

Las cuatro imágenes que se muestran a continuación contienen el procedimiento que deberá aplicarse, sin submuestreo, en este paquete estadístico para construir la base de datos y para generar un análisis de varianza y una comparación de medias entre niveles del Factor A, del Factor B y para cada combinación entre parejas de niveles diferentes de ambos factores, con la prueba de la diferencia mínima significativa de Fisher (DMS o LSD).

InfoStat/L

Archivo Edición Datos Resultados Estadísticas Gráficos Ventanas Aplicaciones Ayuda

CLPD

Caso	H	C	A	B
1	1	1	1	1
2	1	1	1	1
3	1	1	1	1
4	1	1	1	2
5	1	1	1	2
6	1	1	1	2
7	1	1	1	3
8	1	1	1	3
9	1	1	1	3
10	1	2	2	1
11	1	2	2	1
12	1	2	2	1
13	1	2	2	2
14	1	2	2	2

Real Registros: 145*6

- Medidas resumen
- Tablas de frecuencias
- Probabilidades y cuantiles
- Estimación de características poblacionales
- Cálculo del tamaño muestral
- Inferencia basada en una muestra
- Inferencia basada en dos muestras
- Análisis de la varianza
 - Análisis de la varianza no paramétrica
 - Modelos lineales generales y mixtos
 - Modelos lineales generalizados mixtos (MLGM)
- Regresión lineal
- Regresión no lineal
- Modelos no lineales mixtos
- Ridge regression
- Análisis de correlación
- Datos categorizados
- Análisis multivariado
- Serie de tiempo
- Suavizados y ajustes

Max = 1 P05 = 1 P95 = 1

InfoStat/L

Archivo Edición Datos Resultados Estadísticas Gráficos Ventanas Aplicaciones Ayuda

CLPD

Caso	H	C	A	B	S	rg
1	1	1	1	1	1	2.1
2	1	1	1	1	2	1.9
3	1	1	1	1	3	2.3
4	1	1	1	2	1	6.0
5	1	1	1	2	2	5.6
6	1	1	1	2	3	6.5
7	1	1	1	3	1	7.6
8	1	1	1	3	2	7.0
9	1	1	1	3	3	8.0
10	1	2	2	1	1	2.8
11	1	2	2	1	2	2.1
12	1	2	2	1	3	3.0
13	1	2	2	2	1	6.9
14	1	2	2	2	2	6.3

Real Registros: 145*6 n=1 Suma

2(0)

Selección si contiene...

Cancelar Limpiar Aceptar

Análisis de la varianza

Caso S

Variables Particiones ...

Variables dependientes

rg

Variables de clasificación

H C A B

Covariables

InfoStat/L

Archivo Edición Datos Resultados Estadísticas Gráficos Ventanas Aplicaciones Ayuda

CLPD

Caso	H	C	A	B	S	rg
1	1	1	1	1	1	2.1
2	1	1	1	1	2	1.9
3	1	1	1	1	3	2.3
4	1	1	1	2	1	6.0
5	1	1	1	2	2	5.6
6	1	1	1	2	3	6.5
7	1	1	1	3	1	7.6
8	1	1	1	3	2	7.0
9	1	1	1	3	3	8.0
10	1	2	2	1	1	2.8
11	1	2	2	1	2	2.1
12	1	2	2	1	3	3.0
13	1	2	2	2	1	6.9
14	1	2	2	2	2	6.3

Real Registros: 145*6 n=1 Suma

Análisis de la varianza

Modelo Comparaciones Contrastes

Especificación de los términos del modelo

H|H*A
C|H*A
A|H*A
H*A
B
A*B|

Variables de clasificación

H C A B

Covariables

SC Tipo I Sólo generar tabla resumen con

Agregar Interacciones

Guardar

Residuos
 Predichos
 Res. estud.
 Abs(residuos)
 Sobrescribir

Aceptar Cancelar Ayuda

InfoStat/L

Archivo Edición Datos Resultados Estadísticas Gráficos Ventanas Aplicaciones Ayuda

CLPD

Caso	H	C	A	B	S	rg
1	1	1	1	1	1	2.1
2	1	1	1	1	2	1.9
3	1	1	1	1	3	2.3
4	1	1	1	2	1	6.0
5	1	1	1	2	2	5.6
6	1	1	1	2	3	6.5
7	1	1	1	3	1	7.6
8	1	1	1	3	2	7.0
9	1	1	1	3	3	8.0
10	1	2	2	1	1	2.8
11	1	2	2	1	2	2.1
12	1	2	2	1	3	3.0
13	1	2	2	2	1	6.9
14	1	2	2	2	2	6.3

Real Registros: 145*6 n=1 Sum

Análisis de la varianza

Modelo Comparaciones Contrastes

Seleccionar Método de Comparación:

- Ninguna
- LSD Fisher
- DGC
- Jolliffe
- Bonferroni
- Tukey
- Duncan
- SNK
- BSS
- Scott_Knott
- Scheffé

Presentación:

- En lista ascendente
- En lista descendente
- En forma matricial

Nivel significación:

- 0.10
- 0.05
- 0.01
- Otro:

Error

Gráfico de barras

Guardar:

- Residuos
- Predichos
- Res. estud.
- Abs(residuos)
- Sobrescribir

Mostrar medias según:

- H
- C
- A
- H*A
- B
- A*B

SC Tipo I Sólo generar tabla resumen con

Agregar Interacciones

Aceptar Cancelar Ayuda

InfoStat/L - CLPD

Archivo Edición Datos Resultados Estadísticas Gráficos Ventanas Aplicaciones Ayuda

CLPD

Resultados

Análisis de la varianza

Variable	N	R ²	R ² Adj	CV
rg	144	0.98	0.97	6.37

Datos desbalanceados en celdas.
Para otra descomposición de la SC especifique los contrastes apropiados.. !!

Cuadro de Análisis de la Varianza (SC tipo I)

F.V.	SC	gl	CM	F	p-valor (Error)
Modelo.	864.03	23	37.57	228.02	<0.0001
H	1.50	3	0.50	2.74	0.1358 (H*A)
C	0.39	3	0.13	0.72	0.5751 (H*A)
A	90.26	3	30.09	164.85	<0.0001 (H*A)
H*A	1.10	6	0.18	1.11	0.3619
B	769.06	2	384.53	2334.03	<0.0001
A*B	1.72	6	0.29	1.74	0.1168
Error	19.77	120	0.16		
Total	883.80	143			

Test:LSD Fisher Alfa=0.01 DMS=0.37331
Error: 0.1825 gl: 6

A Medias n E.E.

Caso	rg	gl	Medias	Letras
4	7.60	36	0.07	A
3	6.47	36	0.07	B
2	5.94	36	0.07	C
1	5.48	36	0.07	D

Medias con una letra común no son significativamente diferentes (p > 0.01)

Test:LSD Fisher Alfa=0.01 DMS=0.21686
Error: 0.1647 gl: 120

B Medias n E.E.

Caso	rg	gl	Medias	Letras
3	8.67	48	0.06	A
2	7.24	48	0.06	B
1	3.21	48	0.06	C

Medias con una letra común no son significativamente diferentes (p > 0.01)

Test:LSD Fisher Alfa=0.01 DMS=0.43372
Error: 0.1647 gl: 120

A B Medias n E.E.

Caso	rg	gl	Medias	Letras	
4	3	9.98	12	0.12	A
3	3	8.94	12	0.12	B
4	2	8.32	12	0.12	C
2	3	8.07	12	0.12	C D
1	3	7.71	12	0.12	D E
3	2	7.32	12	0.12	E F
2	2	6.94	12	0.12	F
1	2	6.37	12	0.12	G
4	1	4.52	12	0.12	H
3	1	3.16	12	0.12	I
2	1	2.81	12	0.12	I
1	1	2.37	12	0.12	J

Medias con una letra común no son significativamente diferentes (p > 0.01)

ANAVA

J-R Port = disabled/LE

Las cuatro imágenes siguientes corresponden al análisis de los mismos datos cuando se aplica submuestreo balanceado.

InfoStat/L

Archivo Edición Datos Resultados Estadísticas Gráficos Ventanas Aplicaciones Ayuda

CLPD

Caso	H	C	A	B	S	rg
1	1	1	1	1	1	2.1
2	1	1	1	1	2	1.9
3	1	1	1	1	3	2.3
4	1	1	1	2	1	6.0
5	1	1	1	2	2	5.6
6	1	1	1	2	3	6.5
7	1	1	1	3	1	7.6
8	1	1	1	3	2	7.0
9	1	1	1	3	3	8.0
10	1	2	2	1	1	2.8
11	1	2	2	1	2	2.1
12	1	2	2	1	3	3.0
13	1	2	2	2	1	6.9
14	1	2	2	2	2	6.3

Real Registros: 145*6 n=1

Análisis de la varianza

Variables | Particiones ...

Variables dependientes: rg

Variables de clasificación: H, C, A, B, S

Covariables:

Seleccionar si contiene... 1(0)

Cancelar Limpiar Aceptar

InfoStat/L

Archivo Edición Datos Resultados Estadísticas Gráficos Ventanas Aplicaciones Ayuda

CLPD

Caso	H	C	A	B	S	rg
1	1	1	1	1	1	2.1
2	1	1	1	1	2	1.9
3	1	1	1	1	3	2.3
4	1	1	1	2	1	6.0
5	1	1	1	2	2	5.6
6	1	1	1	2	3	6.5
7	1	1	1	3	1	7.6
8	1	1	1	3	2	7.0
9	1	1	1	3	3	8.0
10	1	2	2	1	1	2.8
11	1	2	2	1	2	2.1
12	1	2	2	1	3	3.0
13	1	2	2	2	1	6.9
14	1	2	2	2	2	6.3

Real Registros: 145*6 n=1

Análisis de la varianza

Modelo | Comparaciones | Contrastes

Especificación de los términos del modelo:

- H\H*A
- C\H*A
- A\H*A
- H*A
- B
- A*B
- H*A*B>S

Variables de clasificación: H, C, A, B, S

Covariables:

SC Tipo I Sólo generar tabla resumen con

Agregar Interacciones

Guardar: Residuos, Predichos, Res estud., Abs(residuos), Sobrescribir

Aceptar Cancelar Ayuda

CLPD

Caso	H	C	A	B	S	rg
1	1	1	1	1	1	2.1
2	1	1	1	1	2	1.9
3	1	1	1	1	3	2.3
4	1	1	1	2	1	6.0
5	1	1	1	2	2	5.6
6	1	1	1	2	3	6.5
7	1	1	1	3	1	7.6
8	1	1	1	3	2	7.0
9	1	1	1	3	3	8.0
10	1	2	2	1	1	2.8
11	1	2	2	1	2	2.1
12	1	2	2	1	3	3.0
13	1	2	2	2	1	6.9
14	1	2	2	2	2	6.3

Real Registros: 145*6 n=1 Su

Análisis de la varianza

Modelo Comparaciones Contrastes

Seleccionar Método de Comparación:

Ninguna Duncan
 LSD Fisher SNK
 DGC BSS
 Jolliffe Scott_Knott
 Bonferroni Scheffé
 Tukey

Presentación:

En lista ascendente
 En lista descendente
 En forma matricial

Nivel significación:

0.10 Otro
 0.05
 0.01

Error

Gráfico de barras

SC Tipo I Sólo generar tabla resumen con

Guardar Residuos
 Guardar Predichos
 Guardar Res estud.
 Guardar Abs(residuos)
 Guardar Sobrescribir

CLPD

Resultados

Cuadro de Análisis de la Varianza (SC tipo I)

F.V.	SC	gl	CM	F	p-valor (Error)
Modelo.	878.04	119	7.38	30.73	<0.0001
H	1.50	3	0.50	2.74	0.1358 (H*A)
C	0.39	3	0.13	0.72	0.5751 (H*A)
A	90.26	3	30.09	164.85	<0.0001 (H*A)
H*A	1.10	6	0.18	0.76	0.6082
B	769.06	2	384.53	1601.29	<0.0001
A*B	1.72	6	0.29	1.20	0.3421
H*A*B>S	14.01	96	0.15	0.61	0.9530
Error	5.76	24	0.24		
Total	883.80	143			

Test:LSD Fisher Alfa=0.05 DMS=0.24638
 Error: 0.1825 gl: 6

A Medias n E.E.

Medias	n	E.E.	Letras
4	7.60	36 0.07	A
3	6.47	36 0.07	B
2	5.94	36 0.07	C
1	5.48	36 0.07	D

Medias con una letra común no son significativamente diferentes (p > 0.05)

Test:LSD Fisher Alfa=0.05 DMS=0.20645
 Error: 0.2401 gl: 24

B Medias n E.E.

Medias	n	E.E.	Letras
3	8.67	48 0.08	A
2	7.24	48 0.08	B
1	3.21	48 0.08	C

Medias con una letra común no son significativamente diferentes (p > 0.05)

Test:LSD Fisher Alfa=0.05 DMS=0.41290
 Error: 0.2401 gl: 24

A B Medias n E.E.

Medias	n	E.E.	Letras
4	9.98	12 0.15	A
3	8.94	12 0.15	B
4	8.32	12 0.15	C
2	8.07	12 0.15	D
1	7.71	12 0.15	E
3	7.32	12 0.15	F
2	6.94	12 0.15	G
1	6.37	12 0.15	H
4	4.52	12 0.15	I
3	3.16	12 0.15	J
2	2.81	12 0.15	
1	2.37	12 0.15	

Medias con una letra común no son significativamente diferentes (p > 0.05)

Discusión

Los arreglos en PD y Parcelas Subdivididas para DCA y DBCA, han sido utilizados más frecuentemente en ensayos conducidos en las ciencias agropecuarias y forestales, entre otras (Gomez y Gomez, 1984; Martínez, 1988, 1994; Montgomery, 2009; González *et al.*, 2019), pero la correspondientes al Cuadro Latino (DCL) no está bien documentado (Sahagún, 1998; Ledolter, 2010; Tirado y Tirado, 2017; González *et al.*, 2022). En este estudio se aplica el DCL a los niveles del Factor A en Parcelas Principales, y los correspondientes al Factor B son asignadas a las Subparcelas de manera completamente al azar, como lo sugirieron Smith (1951), Martínez (1994), y Tirado y Tirado (2017). Esto origina dos tipos de error: uno igual al de un DCL para un experimento de un factor, y otro que es el residual en el modelo (Martínez, 1994; <https://biometrics.ilri.org/Publication/Full%20Text/chapter20.pdf>).

En la literatura publicada se ha encontrado poca evidencia con relación al modelo estadístico de referencia, así como para un Análisis de varianza (Anava) y una comparación de medias (Ledolter, 2010; Tirado y Tirado, 2017; <https://biometrics.ilri.org/Publication/Full%20Text/chapter20.pdf>), especialmente para submuestreo (Gomez y Gomez, 1984; Martínez, 1988; PROC ANOVA: Latin Square Split Plot :: SAS/STAT(R) 9.22 User's Guide). Martínez (1994) sólo presentó una tabla con las fuentes de variación (FV) y los grados de libertad (GL) que podrían calcularse sin submuestreo. Gomez y Gomez (1984) mostraron dos tablas con las FV y los GL para un DBCA y para un arreglo en PD con submuestreo balanceado, pero para un solo ensayo; adicionalmente, ellos describieron los procedimientos para obtener un Anava en ambas situaciones, con énfasis en la estimación de los Errores Muestral (EM) y Experimental (EE), que en el presente estudio conforman al Error Conjunto (EC).

En el Sistema para Análisis Estadístico (SAS, Statistical Analysis System) sólo se presentaron los datos proporcionados por Smith (1951; PROC ANOVA: Latin Square Split Plot :: SAS/STAT(R) 9.22 User's Guide) pero sin submuestreo, así como un código para obtener un Anava con tres tipos de errores; con relación a la presente investigación, CxHxA es igual al Error a, y la suma de HxB y del residual del modelo originan al Error b. Tirado y Tirado (2017) presentaron el modelo estadístico para un PD en DCL sin submuestreo, así como un ejemplo para generar su Anava; ellos también generalizaron este tipo de análisis para un arreglo en Parcelas Subdivididas en DCL sin submuestreo.

En el presente estudio se construyó el modelo estadístico para experimentos en PD en DCL, sin y con submuestreo balanceado, y se generaron las fórmulas para calcular GL y suma de cuadrados (SC) aplicando Mínimos Cuadrados y formas cuadráticas o matriciales, con base en las recomendaciones realizadas para otros estudios por González *et al.* (2023) y González *et al.* (2024 a, b), aunque ellos también destacaron el uso de software.

En esta investigación se consideró que no hay interacción entre hileras (H), columnas (C), y niveles del Factor A, como lo sugirieron Gomez y Gomez (1984), Martínez (1988, 1994), Montgomery (2009), y Tirado y Tirado (2017), entre otros, pero aplicando la guía publicada por Sahagún (1998), se llegará a los mismos resultados con las consideraciones siguientes: H anidado en C, o viceversa; H, C o ambos anidados dentro del Factor A; los factores A y B están cruzados; H, C o ambos anidados en la interacción AxB; el error conjunto anidado en todas las componentes de este modelo. Con submuestreo balanceado, además de lo anterior, S estará anidado en H, C y AxB.

Al aplicar InfoGen o InfoStat será correcto elegir cómo Error a, alguna de las siguientes combinaciones: HxC, CxH, HxA, CxA, o HxCxA, debido a que en un DCL también es cierto

que $H=C=A=R$, donde R es el número de repeticiones si se eligiera un DCA o un DBCA. En el presente estudio la interacción $H \times A$ fue considerada como Error a, pero al considerar el código para SAS que permitió analizar los datos de Smith (1951) éste es equivalente a la interacción $C \times H \times A$. Así, las componentes restantes en el modelo sin submuestreo serán B , $A \times B$, y Error b. Con submuestreo balanceado, el EC debe fraccionarse en Errores Muestral (EM) y Experimental (EE), y los componentes restantes serán los mismos que para el PD en DCL sin submuestreo. En ambos paquetes estadísticos EM será estimado directamente, cuando en especificaciones a los términos del modelo se capture la instrucción $H * A * B > S$, pero $C * A * B > S$ originará los mismos resultados.

Al aplicar InfoStat o InfoGen (Balzarini *et al.*, 2008; Di Rienzo *et al.*, 2008; Balzarini y Di Rienzo, 2016), la diferencia que se origine al considerar o no a S , el número de veces que se aplica submuestreo balanceado, permitirá el cálculo indirecto de SC EM, pero para validar resultados, también podrán emplearse las fórmulas presentadas previamente, así como las que fueron construidas para estimar EC y EE, si el usuario usa Mínimos Cuadrados, formulas cuadráticas o matriciales, o ambas. El paquete estadístico STAR genera un plan de aleatorización para PD en DCL, así como un Anava y una comparación de medias para los Factores A , B y para su interacción, pero en este no está implementada la modalidad de submuestreo. Esta misma situación se ha observado al revisar diversos paquetes estadísticos que frecuentemente se utilizan en las ciencias agropecuarias y forestales, entre otras (González *et al.*, 2019; Pérez *et al.*, 2022). También podría construirse una tabla para la interacción $H \times A \times B$; la diferencia entre SC Total y SC de esta triple interacción producirá SC EM. Así, SC EE sería la diferencia entre SC EC y SC EM.

Para un PD en DCL sin submuestreo, las hipótesis estadísticas relacionadas con H, C, y A serán probadas utilizando al Error α , o la interacción HxA, mientras que las correspondientes al Factor B, y a la interacción AxB serán evaluadas con el error conjunto (Ledolter, 2010; Tirado y Tirado, 2017; PROC ANOVA: Latin Square Split Plot :: SAS/STAT(R) 9.22 User's Guide). Para el modelo con submuestreo, primero deberá determinarse si EM es significativo: si lo es, éste será utilizado para probar las hipótesis estadísticas del Factor B y de su interacción; si no lo es, ambas fuentes de variación serán evaluadas utilizando el residual del modelo (Ledolter, 2010; Tirado y Tirado, 2017).

Con este mismo enfoque serán realizadas las pruebas de comparación de medias para las componentes dentro de Parcelas Principales y dentro de Subparcelas (Ledolter, 2010; Tirado y Tirado, 2017), pero también podría recurrirse a las recomendaciones proporcionadas por Gomez y Gomez (1984), Little y Hills (2008), y Sahagún (1998), si es que H, C, pero no ambas, no fueran significativas, en cuyo caso ese tipo de comparaciones serán equivalentes a las de un DBCA en PD. Si H y C tampoco son significativas, el usuario tiene la opción de analizar sus datos como un DCA en PD utilizando la misma base de datos que para los casos referenciados previamente (Balzarini *et al.*, 2008; Balzarini y Di Rienzo, 2016; Di Rienzo *et al.*, 2008; Tirado y Tirado, 2017).

Las metodologías presentadas en esta investigación permitirán homologar fórmulas para calcular grados de libertad y suma de cuadrados de manera fácil y confiable, como lo han mostrado para otros estudios González *et al.* (2023), y González *et al.* (2024 a, b); con su simbología y la descrita en Mendenhall (1987), Sahagún (2007), y Montgomery (2009), se generarán de manera fácil y confiable las formas cuadráticas o matriciales que alimentarán más frecuentemente a las calculadoras de matrices, o a los paquetes estadísticos como SAS,

Agrobase, SPSS, StatGraphics, STAR, PB Tools, entre otros, por lo que los procedimientos para construir un modelo estadístico, para obtener una Anava y para comparar medias de tratamientos en experimentos factoriales, para el tipo de arreglo de unidades experimentales aquí considerado, permitirán verificar directa o indirectamente, los cálculos pertinentes a ambas metodologías de manera más rápida y segura.

Conclusiones

En los dos modelos estadísticos que se construyeron en esta investigación se asumió que no hay interacción entre Hileras, Columnas y niveles del Factor A, asignables a Parcelas Principales en un diseño en Cuadro Latino; adicionalmente se consideró que hay cruzamiento entre los factores A y B. Ambos modelos también se pueden generar si se considera lo siguiente: H anidado en C, o viceversa; H, C o ambos anidados dentro del Factor A; los factores A y B están cruzados; H, C o ambos están anidados en la interacción AxB; el error conjunto se anida en todas las componentes de este modelo. Con submuestreo balanceado, además de lo anterior, S estará anidado en H, C y AxB. Para verificar grados de libertad y sumas de cuadrados en ambos modelos pueden aplicarse las fórmulas alternativas que fueron construidas para los Errores a, Conjunto, Muestral y Experimental, para cada una o para ambas metodologías. Otra opción sería aplicar directamente las fórmulas generadas para Parcelas Principales y Subparcelas, y estimar indirectamente los cuatro errores de referencia.

Literatura citada

Balzarini, M. G. y Di Rienzo, J. A. 2016. InfoGen. FCA. Universidad Nacional de Córdoba, Argentina (<http://www.info-Gen.com.ar>).

Balzarini, M. G.; González, L.; Tablada, M.; Casanoves, F.; Di Rienzo, J. A. y Robledo, C. W. 2008. Manual del usuario de InfoStat. Ed. Brujas, Córdoba, Argentina. 348 pp.

Di Rienzo, J. A.; Casanoves, F.; Balzarini, M. G.; González, L.; Tablada, M.; Robledo, C. W. 2008. InfoStat, versión 2008. Grupo InfoStat, FCA. Universidad Nacional de Córdoba, Argentina. (<https://www.infostat.com.ar>).

Gomez, K. A. y Gomez, A. A. 1984. Statistical procedures for agricultural research. 2nd Ed. John Wiley & Sons, Inc. Printed in Singapore. 680 pp.

González, H. A.; Pérez, L. D. J.; Rubí, A. M.; Gutiérrez, R. F.; Franco, M. J. R. P. y Padilla, L. A. 2019. InfoStat, InfoGen y SAS para contrastes mutuamente ortogonales en experimentos en bloques completos al azar en parcelas subdivididas. *Revista Mexicana de Ciencias Agrícolas*. 10(6):1417-1431.

González, H. A.; Pérez, L. D. J.; Balbuena, M. A.; Franco, M. J. R.; Gutiérrez, R. F. y Rodríguez, G. J. A. 2023. Submuestreo balanceado en experimentos monofactoriales usando InfoStat y InfoGen: validación con SAS. *Revista Mexicana de Ciencias Agrícolas*. 14(2):235-249 (DOI:<https://doi.org/10.29312/remexca.v14i2.3418>).

González, H., A.; Pérez, L., D.J.; Hernández, A., J.; Franco, M., J.R.P.; Rubí, A., M.; Balbuena, M., A. 2024 a. Tratamientos anidados dentro de un arreglo en grupos de bloques completos balanceados. *Revista Mexicana de Ciencias Agrícolas*.15(2), e 3634 (DOI: <https://doi.org/10.29312/remexca.v15i2.3634>).

González, H., A.; Pérez, L., D.J.; Hernández, A., J.; Franco, M., J.R.P.; Balbuena, M., A.; Rubí, A., M. 2024 b. Serie de experimentos para tratamientos anidados en grupos en arreglo de bloques completos balanceados. *Revista Mexicana de Ciencias Agrícolas*. 15(7) (En edición).

Ledolter, J. 2010. Split-plot design: discussion and examples. *International Journal of Quality Engineering and Technology*. 1(4): 441-457 (DOI: <https://doi.org/10.1504/IJQET.2010.035588>).

Little, T. M. y Hills, F. J. 2008. *Métodos Estadísticos para la Investigación en la Agricultura*. Ed. Trillas, SA de CV. México, DF. 270 p. ISBN:978-968-24-3629-1.

Martínez, G. A. 1988. *Diseños experimentales. métodos y elementos de teoría*. Editorial. Trillas, 1ª Ed. México, DF. 756 pp.

Martínez, G., A. 1994. *Experimentación Agrícola: Métodos Estadísticos*. Primera Edición. Universidad Autónoma Chapingo, México, D.F., 357 p (ISBN: 9789 6888 42 867).

Mendenhall, W. 1987. *Introducción a la Probabilidad y la Estadística*. Grupo Editorial Iberoamérica. 1ª Ed. México, DF. 626 pp.

Montgomery, D. C. 2009. *Design and Analysis of Experiments*. 7th Ed. John Wiley & Sons, Inc. U.S.A. 656 pp.

Pérez, L., D.; Jasso, B., G.; Saavedra, G., C.; Franco, M., J.R.P.; Ramírez, D., J. F.; González, H., A. 2022. Uso de artificios en Opstat para analizar series de experimentos en dialélico parcial. *Revista Mexicana de Ciencias Agrícolas*. 13(2) 273-287 (DOI: <https://doi.org/10.29312/remexca.v13i2.3130>).

Sahagún, C. J. 1998. Construcción y análisis de los modelos fijos, aleatorios y mixtos. Departamento de Fitotecnia. Programa Nacional de Investigación en Olericultura. Universidad Autónoma Chapingo (UACH). Boletín Técnico Núm. 2, 64 p.

Sahagún, C. J. 2007. Estadística Descriptiva y Probabilidad: Una Perspectiva Biológica. 2ª Ed. Universidad Autónoma Chapingo (UACH). México, D.F. 282 pp. ISBN: 978 -968-02- 0357-4.

Smith, W., G. 1951. Dissertation notes on Canadian Sugar Factories Ltd, Taber, Alberta, Canadá.

Tirado, E., G., y Tirado, G., D. N. 2017. Tratado de Estadística Experimental. Primera Edición. Editorial Centro de Estudios e Investigaciones para el Desarrollo Docente (CENID, AC), Guadalajara, Jalisco, México (ISBN: 978-607-8435-43-2).

V. CONCLUSIONES GENERALES

Los tres modelos estadísticos que son considerados en esta investigación fueron construidos con la suposición de ausencia de interacción entre Hileras y Columnas, y de éstas con los factores A y B, y con su interacción, para un ensayo, o extendiendo su aplicación a la serie de experimentos cuando se incluyen m ambientes, pero éstos también se pueden generar para llegar a los mismos resultados si se aplica correctamente la guía publicada por Sahagún (1998), haciendo énfasis en los principios de cruzamiento y anidamiento que se pueden definir entre dos o más de los factores en estudio. Las fórmulas para calcular grados de libertad y suma de cuadrados en los ensayos individuales fueron confiables, debido a que los cálculos manuales realizables con las técnicas de mínimos cuadrados y formas cuadráticas o matriciales pueden ser validados con una calculadora de matrices o con InfoStat, InfoGen o cualquier otro paquete estadístico que permita analizar datos sin y con la aplicación de submuestreo balanceado dentro de las unidades o parcelas experimentales. Los procedimientos aritméticos, algebraicos o matriciales que aplican a un PD en DCL, sin y con submuestreo balanceado fueron generalizados a su serie de experimentos, pero solo sin submuestreo, pero a partir de ésta se pueden realizar inferencias hacia las correspondientes al caso de submuestreo balanceado, y su aplicación con los dos paquetes estadísticos previamente referenciados.

VI. BIBLIOGRAFÍA CITADA GENERAL

Cochran, W.G., Cox, G.M. 1978. Diseños experimentales. Editorial Trillas, México, 661 pp.

Fernández, E., R.; Trapero, A.; Domínguez, J. 2010. Experimentación en Agricultura. Ed. Junta de Andalucía. Consejería de Agricultura y Pesca. Serie Agricultura: formación. Sevilla. 350 pp. ISBN 978-84-8474-281-4.

Gavilánez, L., F. 2024. Precisión en los Diseños Agrícolas. 1^{ra} Ed. Ed. Mawil Publicaciones de Ecuador. Quito, Ecuador. 329 pp. (DOI; <https://doi.org/10.26820/978-9942-654-06-9>).

Gomez, K. A. y Gomez, A. A. 1984. Statistical procedures for agricultural research. 2nd Ed. John Wiley & Sons, Inc. Printed in Singapore. 680 pp.

González, H. A.; Pérez, L. D. J.; Balbuena, M. A.; Franco, M. J. R.; Gutiérrez, R. F. y Rodríguez, G. J. A. 2023. Submuestreo balanceado en experimentos monofactoriales usando InfoStat y InfoGen: validación con SAS. Revista Mexicana de Ciencias Agrícolas. 14(2):235-249. DOI:<https://doi.org/10.29312/remexca.v14i2.3418>.

Little, T. M. y Hills, F. J. 2008. Métodos Estadísticos para la Investigación en la Agricultura. Ed. Trillas, SA de CV. México, DF. 270 pp. ISBN:978-968-24-3629-1.

Martínez, G., A. 1994. Experimentación Agrícola: Métodos Estadísticos. Primera Edición. Universidad Autónoma Chapingo, México, D.F., 357 pp (ISBN: 9789 6888 42 867).

Montgomery, D., C. 2004. Diseño y Análisis de Experimentos. 2^{da} Ed. Ed. Limusa, S.A. de C.V. Grupo Noriega Editores. México, D.F. 686 pp. ISBN 968-18-6156-6.

Ojeda, R., M., M.; Gutiérrez, R., G. 2005. Propuesta de un Modelo Lineal Jerárquico para Analizar una Serie de Experimentos. Revista Agrociencia, vol. 39, núm. 5, septiembre-octubre. pp. 529-538. ISSN: 1405-3195. <https://www.redalyc.org/articulo.oa?id=30239506>.

Pantoja C. 1989. Algunos criterios y procedimientos estadísticos en la experimentación en fincas de agricultores. Ramakrishna B, II curso corto. Pruebas en fincas (pp. 218 – 236). Quito: Programa Cooperativo de Investigación Agrícola para la Subregión Andina – PROCIANDINO.

Zamudio, S., F., J. y Alvarado, S., A., A. 1996. Análisis de Diseños Experimentales con Igual Número de Submuestras. Ed. Universidad Autónoma Chapingo. División de Ciencias Forestales. Chapingo, México. 58 pp.