

UNIVERSIDAD AUTÓNOMA DE CHIHUAHUA

FACULTAD DE ZOOTECNIA Y ECOLOGÍA

SECRETARÍA DE INVESTIGACIÓN Y POSGRADO



**PRUEBA DE IGUALDAD DE UN CONJUNTO DE MODELOS
LINEALES Y SUS APLICACIONES**

POR:

M. S. M. ARWELL NATHÁN LEYVA CHÁVEZ

**TESINA PRESENTADA COMO REQUISITO PARA OBTENER EL
GRADO DE MAESTRÍA PROFESIONAL EN ESTADÍSTICA
APLICADA**



Prueba de igualdad de un conjunto de modelos lineales. Tesina presentada por Arwell Nathán Leyva Chávez como requisito parcial para obtener el grado de Maestría Profesional en Estadística Aplicada, ha sido aprobado y aceptada por:

Ph.D. Carlos Ortega Ochoa
Director de la Facultad de Zootecnia y Ecología

D.Ph. Agustín Corral Luna
Secretario de Investigación y Posgrado

Ph.D Ana Luisa Rentería Monterrubio
Coordinador Académico

Dr. José Antonio Díaz García
Presidente

19 DICIEMBRE 2018

Fecha

Comité:
M.C. Oscar Alejandro Martínez
Jaime
Ph.D. Felipe Alonso Rodríguez
Almeida
Dr. Nicolás Callejas Juárez

© Derechos Reservados
AUTOR. ARWELL
NATHÁN LEYVA CHÁVEZ
DIRECCIÓN: PERIFÉRICO
FRANCISCO R. ALMADA
KM. 1, CHIHUAHUA,
CHIH., MÉXICO C.P. 31453
DICIEMBRE 2018

AGRADECIMIENTOS

Agradezco primero y antes que nada a mi familia, a mi esposa por todo el tiempo invertido en este proyecto, ya que sin sus ánimos y comprensión esto no sería posible; a mis hijos Arwell y Axel, que les tuve que decir varias veces que tenía que trabajar y lo aceptaron sin reclamos, tiempo que pienso en su momento no recuperar, pero si pasarlo con calidad; a mi madre por haberme forjado como la persona que soy en la actualidad.

Mi más sincero agradecimiento al Doctor José Antonio Díaz García, director de esta tesina, ya que sin su disponibilidad, paciencia, criterio y tiempo invertido no hubiera podido realizarla. Su apoyo incondicional y, sobre todo, esa gran amistad que me brindó y que me brinda, por escucharme y aconsejarme siempre, ha sido un aporte invaluable.

A mi co-director de tesina y mis revisores por estar siempre al pendiente del proceso del proyecto y, más que nada, en la mejor disposición para que se lograra.

A mis maestros de la Facultad de Zootecnia y Ecología, que directa o indirectamente me ayudaron a la culminación de esta tesina de maestría, por su atención y amabilidad en todo lo referente a mi vida como alumno de la maestría.

A mis amigos y compañeros que estuvieron al pendiente de la culminación de este proceso, siempre con una palabra de apoyo.

Al Sindicato del Personal Académico de la Universidad Autónoma de Chihuahua por el apoyo y la beca otorgada para cursar la maestría.

DEDICATORIA

A Dios, por haberme permitido llegar hasta este punto y haberme dado salud para lograrlo.

A mi familia, por ser la razón de mí existir. A mi esposa, quien tuvo que correr (como dice ella) durante todo este tiempo, por creer siempre en mí. Aunque hemos pasado momentos difíciles, siempre has estado apoyándome y brindándome todo tu amor, por eso, de todo corazón te lo agradezco.

A mi hijo Arwell, que aunque llegara tarde a casa, me esperaba con ansias para jugar.

Con todo mi amor y cariño a mi hijo Axel, quien ya nos espera en el cielo.

A todos aquellos que con sus palabras de aliento me ayudaron a dar este paso en vida.

CURRICULUM VITAE

El autor nació el 3 de julio de 1981 en la ciudad de Chihuahua, Chihuahua, México.

- | | |
|------------------|---|
| 1998-2003 | Licenciatura en el Instituto Tecnológico de Chihuahua II |
| 2004-2011 | Maestría en Sistemas de Manufactura en el Instituto Tecnológico de Chihuahua |
| 2004- a la fecha | Docente del Colegio de Estudios Científicos y Tecnológicos del Estado de Chihuahua Plantel #6 |
| 2005-2011 | Docente del Colegio de Bachilleres de Chihuahua Plantel #10 |
| 2007-2009 | Ingeniero de proyectos en la empresa SAFRAN GROUP LABINAL |
| 2010- a la fecha | Profesor-Investigador de la Facultad de Ciencias Agrotecnológicas de la Universidad Autónoma de Chihuahua |

Ponente en cursos relacionados con la docencia (diseño de programas analíticos, portafolio de evidencias, proyectos formativos, modelo educativo).

Ponente en cursos relacionados con probabilidad y estadística aplicada.

Ponente en diplomado de PROFORDEMS para docentes de nivel medio superior.

RESUMEN

PRUEBA DE IGUALDAD DE UN CONJUNTO DE MODELOS LINEALES Y SUS APLICACIONES

POR:

M.S.M. ARWELL NATHÁN LEYVA CHÁVEZ

Maestría Profesional en Estadística Aplicada

Secretaría de Investigación y Posgrado

Facultad de Zootecnia y Ecología

Universidad Autónoma de Chihuahua

Presidente: Dr. José Antonio Díaz García

Motivada en los resultados de ensayos observacionales o experimentales, uno de los objetivos de la estadística es llevar a cabo comparaciones de parámetros, estructuras y modelos. En este trabajo se estudió la comparación de tres modelos linealizables que tienen la misma forma funcional, para los cuales, las variables dependientes e independientes que intervienen en tales modelos pueden ser de diferente naturaleza (por ejemplo, diferentes escalas), y sus correspondientes tamaños de muestras también pueden ser diferentes. Para este problema es obtenido un estadístico de prueba (W) que permite tomar decisiones sobre la igualdad de un conjunto de modelos linealizables. Basados en la teoría del modelo lineal general y, en particular en el estudio de la igualdad de la intersección, paralelismo y misma ordenada al origen de un conjunto de líneas, tal comparación sobre la igualdad de un conjunto de modelos linealizables es estudiada. Como resultado, un estadístico de prueba con distribución F-Snedecor es obtenido y la correspondiente prueba de hipótesis con su regla de decisión es planteada. Adicionalmente, una base de datos reales en el contexto del control integrado de

plagas es analizada bajo la prueba de hipótesis obtenida. Como se desprende de esta tesina, reescribiendo los modelos linealizables como un modelo lineal general es posible proponer una prueba de hipótesis, de tamaño predefinido, para decidir sobre la igualdad de un conjunto de modelos.

ABSTRACT

PROOF OF EQUALITY OF A SET OF LINEAR MODELS AND THEIR APPLICATIONS

BY:

M.S.M. ARWELL NATHÁN LEYVA CHÁVEZ

Motivated in the results of observational or experimental tests, one of the objectives of statistics is to carry out comparisons of parameters, structures and models. In this paper, the comparison of several linearizable models that have the same functional form is studied, for which the dependent and independent variables that intervene in such models can be of different nature (for example, different scales), and their corresponding sample sizes, can also be different. For this problem, a statistical test is obtained which allows to make decisions about the equality of a set of linear models. Based on the theory of the general linear model and, in particular in the study of the equality of the intersection, parallelism and same ordered to the origin of a set of lines, such a comparison on the equality of a set of linearizable models is studied. As a result, a statistical test with F-Snedecor distribution is obtained and the corresponding hypothesis test with its decision rule is proposed. Additionally, a real database in the context of integrated pest control is analyzed under the hypothesis test obtained. As can be seen from this thesis, by rewriting linearizable models as a general linear model, it is possible to propose an hypothesis test, of predefined size, to decide on the equality of a set of models.

CONTENIDO

RESUMEN	vi
ABSTRACT	viii
CONTENIDO	ix
LISTA DE CUADROS	xi
LISTA DE GRÁFICAS	xii
LISTA DE FIGURAS.....	xiii
LISTA DE FIGURAS DEL APÉNDICE	xiv
INTRODUCCIÓN.....	1
REVISIÓN DE LITERATURA.....	3
MATERIALES Y MÉTODOS	5
Descripción de la Situación a Abordar	5
Metodología para la Propuesta Estadística	8
Metodología para la Aplicación.....	10
RESULTADOS Y DISCUSIÓN	13
Paralelismo, Intersección e Igualdad de Modelos	13
Prueba para la Igualdad de un Conjunto de Modelos Lineales.....	24
Aplicación	31
CONCLUSIONES Y RECOMENDACIONES	45
LITERATURA CITADA	46
APÉNDICE A	48
Elementos del Álgebra de Matrices.....	48
Propiedades de la transpuesta y la inversa.....	49
Formas cuadráticas.	51
Eigenvalores y eigenvectores.	53
Matrices particionadas.....	55
Matrices no negativas.....	55
Inversa generalizada o inversa de Moore-Penrose.....	57
Inverso condicional.....	58
Sistema de ecuaciones lineales.....	59
APÉNDICE B	61

Distribución Normal Multivariada.....	61
Distribución de Formas Cuadráticas	64
Distribución Ji-Cuadrada.....	64
Distribución Ji-Cuadrada no centrada.....	65
Independencia de formas lineales y cuadráticas.....	66
Distribución F no centrada.....	67
APÉNDICE C.....	69
Modelo Lineal General.....	69
Estimación de formas lineales.....	75
Prueba de una hipótesis lineal.....	77
Invariabilidad de la prueba de hipótesis $H\beta = h$	80
Intervalos de confianza y pruebas de hipótesis.....	83
Prueba de la hipótesis $H\beta = h$ usando el método de la razón de verosimilitud	86
APÉNDICE D.....	92
Secuencia para comparación de modelos en el software R	92

LISTA DE CUADROS

Cuadro		Página
1	Análisis de varianza de <i>Diabrotica undecimpunctata</i> en términos de temperatura (x_1) y precipitación (x_2).....	38
2	Análisis de varianza de <i>Diabrotica balteata</i> en términos de temperatura (x_1) y precipitación (x_2).....	39
3	Análisis de varianza de <i>Diabrotica virgifera</i> en términos de temperatura (x_1) y precipitación (x_2).....	40
4	Valores de la prueba de F, probabilidad P y coeficientes de determinación parciales R^2	41
5	Puntos críticos de temperatura y precipitación (x_1, x_2) y valores máximos del número de adultos de tres especies de <i>Diabrotica</i> (Y).....	43

LISTA DE GRÁFICAS

Gráfica		Página
1	Curva de crecimiento estimada para <i>Diabrotica undecimpunctata</i>	34
2	Curva de crecimiento estimada para <i>Diabrotica balteata</i>	35
3	Curva de crecimiento estimada para <i>Diabrotica virgifera</i>	36

LISTA DE FIGURAS

Figura		Página
1	Muestra tomada de una población.....	6
2	Muestra tomada de diferentes poblaciones.....	9

LISTA DE FIGURAS DEL APÉNDICE

Figura		Página
3	Distribución $F_{q, n - q}(\alpha)$ para probar $H\beta = h$	81

INTRODUCCIÓN

En el análisis estadístico de la información de ensayos experimentales, observacionales o fenómenos en general, en donde se pretenden hacer comparaciones sobre el efecto de R tratamientos o el comportamiento de R especies animales, vegetales u entes inertes, basados en la variable de decisión Y (variable dependiente), es usual concluir en términos de la comparación de las R medias aritméticas o, en el mejor de los casos, se realiza una comparación simultánea de las R medias y sus correspondientes varianzas (de la variable Y). Sin embargo, cuando dicho comportamiento cambia a través del tiempo, del espacio, o de variables que son función del tiempo o el espacio -denotemos esta variables como X_1, \dots, X_p (variables independientes)- en ciertas ocasiones es conveniente proponer modelos que describan el comportamiento de las variables de interés $Y_\rho = f_\rho(X_1, \dots, X_p), \rho = 1, \dots, R$, en términos de dichas variables independientes. Entonces, alternativamente a comparar las R medias y/o varianzas de las medidas repetidas o independientes, se puede pensar en comparar los R modelos $Y_\rho = f_\rho(X_1, \dots, X_p)$.

Adicionalmente, una vez estimados estos modelos basados en la información del experimento a través de las técnicas del modelo lineal general, permiten estudiar el comportamiento de las variables dependientes en todo el recorrido de las variables independientes, característica que facilita el estudio de tendencias y hacer predicciones para valores de interés de las variables independientes. Además, es posible estimar los puntos críticos (valores de las variables independientes) y los correspondientes óptimos de las variables dependientes.

Usando modelos de regresión simple y modelos de regresión lineal, esta práctica de comparar R modelos es común en estudios de ensayos médicos, farmacológicos y, más recientemente, en el análisis de experimentos agrícolas.

El presente trabajo tiene como objetivo principal proponer una prueba estadística para la hipótesis de igualdad de R modelos lineales o linealizables, basada en la teoría del modelo lineal general bajo normalidad.

REVISIÓN DE LITERATURA

En primera instancia, se hizo una breve recopilación de notación y tópicos del álgebra de matrices, los cuales son necesarios para la comprensión del resto del material expuesto, resaltando algunos resultados sobre el inverso de matrices particionadas y el inverso de Moore-Penrose (Graybill, 1976; ver Apéndice A).

Asimismo, se estudiaron la distribución normal multivariada y las distribuciones asociadas a las formas cuadráticas, como la Ji-Cuadrada y F-Snedecor centradas y no centradas (ver Apéndice B).

Se analizó el modelo lineal general; en particular se estudió el problema de estimación, propiedades de los estimadores y, en especial, la prueba de hipótesis lineal general (ver Apéndice C). Para concluir esta sección, se estableció el estadístico de prueba para la hipótesis lineal general a través del criterio de la razón de verosimilitud vía la técnica de optimización de multiplicadores de Lagrange (Graybill, 1976).

La prueba de igualdad de modelos linealizables con la misma forma funcional es poco conocida en la literatura. Draper y Smith (1981) hacen un breve comentario sobre el posible interés de estudiar este problema. Por otro lado, Graybill (1976) estudió en detalle este problema en el caso univariado, proponiendo una prueba de hipótesis de tamaño α , con su correspondiente estadístico de prueba y la regla de decisión. Con un enfoque bayesiano, Novick *et al.* (2013) determinaron si las respuestas biológicas bajo la aplicación de dos sustancias son similares. Para ello, dos modelos de regresión lineal simple fueron ajustados y concluyeron haciendo una prueba de paralelismo. También, bajo modelos de regresión no lineales,

específicamente modelos tipo logístico, determinaron si dos ambientes biológicos diferentes darían curvas de dosis-respuesta similares a la misma sustancia, proponiendo una prueba de paralelismo para este tipo de modelos logísticos. Fleetwood *et al.* (2015) llevaron a cabo un estudio paralelo a los autores anteriores, pero ellos desde un punto de vista estadístico clásico. El estudio de paralelismo ha sido de interés desde un punto de vista multivariado. Así, bajo estadística clásica, Díaz y Caro (2006) estudiaron el problema de la igualdad de un conjunto de modelos lineales multivariados linealizables. Desde un punto de vista práctico, la igualdad de modelos lineales multivariados linealizables y su correspondiente optimización multiobjetivo fue usada por Martínez *et al.* (2014) en un problema sobre el control integrado de plagas en el estado de Guanajuato.

MATERIALES Y MÉTODOS

Descripción de la Situación a Abordar

En diferentes áreas del conocimiento, es común estar interesados en describir el comportamiento de ciertas variables dependientes en término de otro conjunto de variables independientes a través de un modelo matemático. En muchas de estas situaciones, no se conoce de antemano la forma funcional de dicho modelo, pero se tiene una muestra de las variables de interés, lo cual permite proponer diferentes formas funcionales del modelo y estimarlas a través de técnicas estadísticas. Por ejemplo, dado un conjunto de datos, se desea saber cuál distribución probabilística sigue la muestra. Alternativamente, supóngase que se tiene una muestra de la variable Y (dependiente) y la variable independiente X . El objetivo es proponer el modelo matemático que mejor describa el comportamiento de la variable Y en términos de la variable X , es decir, se desea encontrar un modelo de la forma

$$Y = f(x),$$

para alguna función $f(\cdot)$. En este caso, por información previa del fenómeno a ser modelado o conociendo la forma del modelo propuesto en casos de variables con comportamientos similares, se proponen varias formas funcionales para el modelo $f(\cdot)$ y se procede a estimar los parámetros que intervienen en dichos modelos como se muestra en la Figura 1. La pregunta ahora es: ¿Cómo decidir cuál es el modelo que mejor ajusta a las observaciones?



Figura 1. Muestra tomada de una población.

Basados en el factor de Bayes, en la teoría de codificación y en la teoría de la información, diferentes criterios de selección de modelos se han propuesto en la literatura estadística para responder a esta pregunta. Estos criterios, como el criterio de Schwarz (SBC), criterio de información de Akaike (AIK), criterio de información de Kullback-Leibler (KLIC), etc., son tan generales que permiten estimar la dimensión del modelo, es decir, la dimensión del espacio paramétrico que mejor describe el comportamiento de la variable Y en función de la variable X (Schwarz, 1978; Kass y Raftery, 1995).

En el presente proyecto el objetivo principal es estudiar un problema alternativo. En este caso se está interesado en comparar $\rho, \rho = 1, 2, \dots, R$, modelos que funcionalmente son iguales (es decir, tienen la misma forma funcional) y pueden ser estimados a través de la teoría del modelo lineal general.

Por ejemplo, si Y denota la variable dependiente y X_1, X_2 son las variables independientes, cada modelo ρ puede ser de la forma:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2$$

$$Y = \exp\{\beta_0 + \beta_1 x_1 + \beta_2 x_2\}$$

$$Y = \exp\{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2\}$$

$$Y = \beta_0 x_1^{\beta_1} x_2^{\beta_2},$$

entre otras muchas posibilidades (Graybill, 1976; Draper y Smith, 1981).

Note que esta comparación de modelos difiere de la descrita anteriormente en diferentes aspectos:

1. Las muestras provienen de diferentes poblaciones o sub-poblaciones como lo muestra la Figura 2.
2. Las variables dependientes Y 's pueden ser diferentes entre cada muestra.
3. El número de variables independientes X 's que intervienen en cada modelo es el mismo, aunque se pueden definir diferentes variables entre muestras.
4. El tamaño de cada muestra puede ser diferente.

Esto es, el número de variables independientes p es el mismo para todos los modelos, pero cabe la posibilidad de que

$$X_1 \neq X_2 \neq \dots \neq X_R$$

Más aún, observe que el tamaño de muestra para cada $p, p = 1, 2, \dots, R$, admite la posibilidad de que sean diferentes, esto es, $n_1 \neq n_2 \neq \dots \neq n_R$, el tamaño de muestra para cada modelo es diferente, como se muestra a continuación,

$$X_1 = \begin{bmatrix} 1 & 3 \\ 1 & 7 \\ 1 & 9 \\ 1 & 8 \end{bmatrix}, X_2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix}, X_3 = \begin{bmatrix} 1 & 6.8 \\ 1 & 3.2 \\ 1 & 9.3 \\ 1 & 5.4 \\ 1 & 1.3 \\ 1 & 7.4 \\ 1 & 8.5 \end{bmatrix}$$

Metodología para la Propuesta Estadística

Con el propósito de establecer el estadístico para la prueba de la igualdad de un conjunto de R modelos, se examinaron las siguientes hipótesis de paralelismo, intersección e igualdad de ordenada al origen de modelos lineales simples; esto es, se establecieron los estadísticos de prueba para las hipótesis:

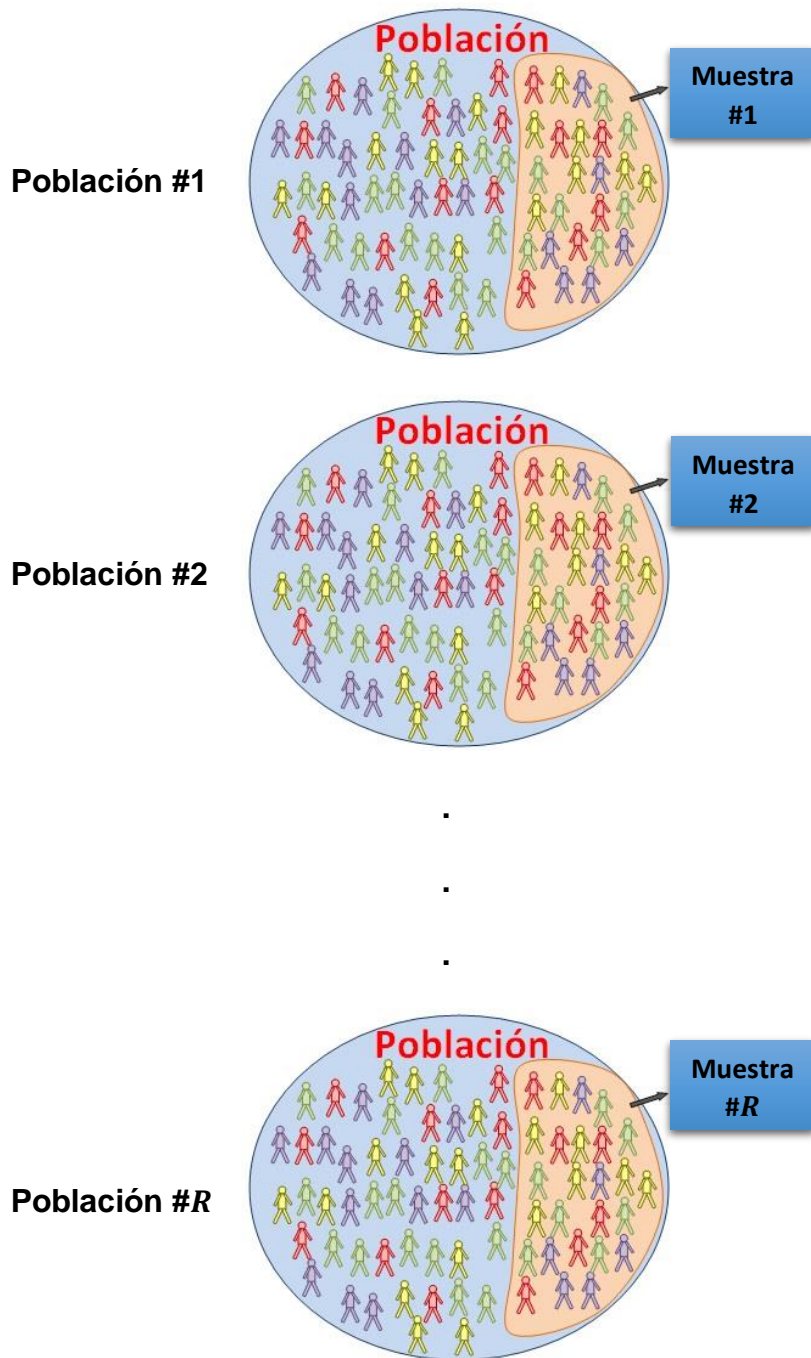


Figura 2. Muestra tomada de diferentes poblaciones.

- 1) Los R modelos lineales simples son paralelos.
- 2) Los R modelos lineales simples tienen la misma ordenada al origen.
- 3) Los R modelos se interseccionan exactamente en un punto x_0 específico.

Con el estudio de estas tres pruebas, se establecen los preliminares básicos para poder estudiar el objetivo fundamental de la presente tesis, la prueba de hipótesis para el problema de comparación de modelos con formas funcionales iguales (en conjunto). Así, en la segunda sección, se deriva el estadístico de prueba para la hipótesis de que R modelos lineales generales son idénticos.

Metodología para la Aplicación

Típicamente, la comparación entre poblaciones de plagas se realiza comparando sus medias, ya sea en medidas independientes o en medidas repetidas, según el caso (Pérez *et al.*, 2010; Murtaugh *et al.*, 2012). Sin embargo, se observó que al llevar al campo las conclusiones así obtenidas, el manejo integrado de plagas no era eficiente. Luego, se ajustaron modelos de regresión polinomiales para describir el comportamiento de la población de plagas en términos de variables climáticas, principalmente temperatura y precipitación (Hodgson *et al.*, 2011; Quesada *et al.*, 2012). Sin embargo, la comparación de las poblaciones de plagas se seguía haciendo en términos de las medias. Recientemente, Martínez-Jaime *et al.* (2014) proponen como alternativa que la dinámica de las poblaciones de plaga sea comparada mediante sus correspondientes modelos de curvas de crecimiento estimadas a través de los modelos de regresión polinomial.

A través de los siguientes ejemplos se ilustrará la utilización de las pruebas de paralelismo, intersección e igualdad de modelos.

Ejemplo 1. Un fertilizante comercial es aplicado al cultivo de trigo para determinar cómo afecta su rendimiento. Se estudian tres variedades de trigo y, para cada variedad, la relación entre el rendimiento en costales y la cantidad de fertilizante aplicado en Kg/Ha x está dada por

$$y_{1j} = \alpha_1 + \beta_1 x_{1j} + \varepsilon_{1j}, \quad y_{2j} = \alpha_2 + \beta_2 x_{2j} + \varepsilon_{2j}, \quad y_{3j} = \alpha_3 + \beta_3 x_{3j} + \varepsilon_{3j}.$$

El investigador está interesado en determinar si las líneas son paralelas, esto es,

$$\beta_1 = \beta_2 = \beta_3,$$

si el incremento en rendimiento promedio por unidad de fertilizante es el mismo para las tres variedades de trigo. O quizás desea determinar si los interceptos son iguales, esto es, si

$$\alpha_1 = \alpha_2 = \alpha_3,$$

(si el rendimiento de cada variedad de trigo es el mismo cuando no se ha aplicado el fertilizante).

Ejemplo 2. Un estudio clínico es llevado a cabo para determinar la relación entre la edad x y la tensión arterial y de individuos entre 50 y 75 años de edad. Para tal efecto se ha asumido un modelo lineal simple, pero el estudio contempla ambos géneros, masculino y femenino. Por lo tanto, se ha decidido que un modelo para cada género es más apropiado. Los modelos son:

$$y_{1j} = \alpha_1 + \beta_1 x_{1j} + \varepsilon_{1j} \text{ para mujeres, y}$$

$$y_{2j} = \alpha_2 + \beta_2 x_{2j} + \varepsilon_{2j} \text{ para los hombres.}$$

Enseguida, ya establecidos los preliminares básicos, esta última técnica (comparación de modelos con formas funcionales iguales en conjunto) se aplicó en la comparación de las curvas de crecimiento de las poblaciones de adultos de *Diabrotica undecimpunctata* (Harold), *Diabrotica balteata* (LeConte) y *Diabrotica virgifera* (Krysan & Smith) en Irapuato, Guanajuato, México.

Utilizando las técnicas de regresión múltiple univariadas se estimaron las curvas de crecimiento poblacional mediante polinomios de tercer grado en términos de la temperatura y la precipitación. Luego, alternativamente a la comparación de las medias de las tres especies en diferentes tiempos, los modelos de dichas curvas de crecimiento polinómicas fueron comparados entre sí.

RESULTADOS Y DISCUSIÓN

Paralelismo, Intersección e Igualdad de Modelos

Algunos problemas de gran interés desde un punto de vista aplicado son los siguientes: determinar si en un conjunto de líneas todas se interseccionan en un punto específico, todas son paralelas o todas son idénticas.

El interés del investigador fue estudiar los modelos por separado y simultáneamente. Al analizar los dos modelos juntos, el interés se centró en probar simultáneamente que

$$\alpha_1 = \alpha_2$$

$$\beta_1 = \beta_2.$$

Esto es, ¿los modelos para ambos géneros son iguales?

En la primera sección son derivados los estadísticos de prueba para las siguientes hipótesis:

- 1) R modelos lineales simples son paralelos.
- 2) R modelos lineales simples tienen el mismo intercepto.
- 3) Los R modelos se interseccionan exactamente en un punto x_0 específico.

En la segunda sección, se deriva el estadístico de prueba para la hipótesis de que R modelos lineales generales son idénticos.

1. Pruebas de hipótesis acerca de R modelos lineales simples.

Primero se discuten algunos detalles sobre el modelo lineal simple y se introduce alguna notación.

Considere los siguientes R modelos lineales simples

$$\begin{aligned}
 y_{1j} &= \alpha_1 + \beta_1 x_{1j} + \varepsilon_{1j} , & j &= 1, 2, \dots, n_1; \\
 y_{2j} &= \alpha_2 + \beta_2 x_{2j} + \varepsilon_{2j} , & j &= 1, 2, \dots, n_2; \\
 y_{3j} &= \alpha_3 + \beta_3 x_{3j} + \varepsilon_{3j} , & j &= 1, 2, \dots, n_3; \\
 &\vdots & & \\
 &\vdots & & \\
 &\vdots & & \\
 y_{Rj} &= \alpha_R + \beta_R x_{Rj} + \varepsilon_{Rj} , & j &= 1, 2, \dots, n_R.
 \end{aligned} \tag{1}$$

Donde $n_1 + n_2 + \dots + n_R = N$, $n_\rho > 2$, $\rho = 1, \dots, R$ y $\varepsilon_{ij} \sim N(0, \sigma^2)$

independientes (ver Apéndice B).

Las hipótesis de interés son:

- i. $H_0: \beta_1 = \beta_2 = \dots = \beta_R$, esto es, las líneas son paralelas.
- ii. $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_R$, esto es, las líneas tienen un intercepto en común.
- iii. $H_0: \alpha_1 + \beta_1 x_0 = \alpha_2 + \beta_2 x_0 = \dots = \alpha_R + \beta_R x_0$ (x_0 conocido), esto es, las líneas se interseccionan en el punto especificado x_0 .

Con el fin de establecer los correspondientes estadísticos de prueba, obsérvese que los R modelos lineales simples se pueden reescribir como un modelo lineal general $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ (ver Apéndice C). Pero antes nótese que el ρ -ésimo modelo lineal simple se puede escribir como

$$\mathbf{Y}_\rho = \mathbf{X}_\rho \boldsymbol{\beta}_\rho + \boldsymbol{\varepsilon}_\rho.$$

Donde

$$\mathbf{y}_\rho = \begin{bmatrix} y_{\rho_1} \\ y_{\rho_2} \\ \vdots \\ y_{\rho_{n_R}} \end{bmatrix}, \quad \mathbf{X}_\rho = \begin{bmatrix} 1 & x_{\rho_1} \\ 1 & x_{\rho_2} \\ \vdots & \vdots \\ 1 & x_{\rho_{n_R}} \end{bmatrix}, \quad \boldsymbol{\beta}_\rho = \begin{bmatrix} \alpha_\rho \\ \beta_\rho \end{bmatrix}, \quad \boldsymbol{\varepsilon}_\rho = \begin{bmatrix} \varepsilon_{\rho_1} \\ \varepsilon_{\rho_2} \\ \vdots \\ \varepsilon_{\rho_{n_R}} \end{bmatrix}.$$

Así los R modelos se pueden reescribir como

$$\mathbb{Y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_R \end{bmatrix}, \quad \mathbb{X} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_R \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_R \end{bmatrix}, \quad \mathbb{E} = \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_R \end{bmatrix}.$$

Obsérvese que, dado que $\varepsilon_{ij} \sim N(0, \sigma^2)$ y son independientes

$$\boldsymbol{\varepsilon}_\rho \sim N_{N_R}(\mathbf{0}, \sigma^2 \mathbf{I}_{n_\rho}).$$

Luego

$$\mathbb{E} \sim N_N(\mathbf{0}, \sigma^2 \mathbf{I}_N).$$

Además,

$$\mathbb{X}'\mathbb{X} = \begin{bmatrix} \mathbf{X}_1' & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2' & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_R' \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_R \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1' \mathbf{X}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2' \mathbf{X}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_R' \mathbf{X}_R \end{bmatrix},$$

y

$$\mathbb{X}'\mathbb{Y} = \begin{bmatrix} \mathbf{X}_1' & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2' & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_R' \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_R \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1' \mathbf{y}_1 \\ \mathbf{X}_2' \mathbf{y}_2 \\ \vdots \\ \mathbf{X}_R' \mathbf{y}_R \end{bmatrix}.$$

De donde

$$(\mathbb{X}'\mathbb{X})^{-1} = \begin{bmatrix} (\mathbf{X}'_1\mathbf{X}_1)^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & (\mathbf{X}'_2\mathbf{X}_2)^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & (\mathbf{X}'_R\mathbf{X}_R)^{-1} \end{bmatrix}.$$

Por lo tanto,

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= (\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}'\mathbb{Y} = \begin{bmatrix} (\mathbf{X}'_1\mathbf{X}_1)^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & (\mathbf{X}'_2\mathbf{X}_2)^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & (\mathbf{X}'_R\mathbf{X}_R)^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{X}'_1\mathbf{y}_1 \\ \mathbf{X}'_2\mathbf{y}_2 \\ \vdots \\ \mathbf{X}'_R\mathbf{y}_R \end{bmatrix} \\ &= \begin{bmatrix} (\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1\mathbf{y}_1 \\ (\mathbf{X}'_2\mathbf{X}_2)^{-1}\mathbf{X}'_2\mathbf{y}_2 \\ \vdots \\ (\mathbf{X}'_R\mathbf{X}_R)^{-1}\mathbf{X}'_R\mathbf{y}_R \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1^-\mathbf{y}_1 \\ \mathbf{X}_2^-\mathbf{y}_2 \\ \vdots \\ \mathbf{X}_R^-\mathbf{y}_R \end{bmatrix}. \end{aligned}$$

Esto es, $\hat{\boldsymbol{\beta}}_\rho = \mathbf{X}_\rho^-\mathbf{y}_\rho$ también.

$$\begin{aligned} (N - 2R)\hat{\sigma}^2 &= \|\mathbb{Y} - \mathbb{X}\hat{\boldsymbol{\beta}}\|^2 = \mathbb{Y}'\mathbb{Y} - \hat{\boldsymbol{\beta}}'\mathbb{X}'\mathbb{Y} = \sum_{\rho=1}^R \mathbf{y}_\rho'\mathbf{y}_\rho - \sum_{\rho=1}^R \mathbf{y}_\rho'\mathbf{X}_\rho\mathbf{X}_\rho^-\mathbf{y}_\rho \\ &= \sum_{\rho=1}^R \mathbf{y}_\rho'(I - \mathbf{X}_\rho\mathbf{X}_\rho^-)\mathbf{y}_\rho. \end{aligned}$$

Esencialmente estos resultados establecen que $\hat{\alpha}_\rho$ y $\hat{\boldsymbol{\beta}}_\rho$ son calculados a partir de los datos del ρ -ésimo modelo y que $\hat{\sigma}^2$ es calculada como una ponderación de los estimadores de $\hat{\sigma}^2$ para cada modelo. Adicionalmente, $\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \dots, \hat{\boldsymbol{\beta}}_R, \hat{\sigma}^2$ son estimadores suficientes.

Para el modelo lineal simple recuérdese que

$$y_i = \alpha + \beta x_i + \varepsilon_i.$$

Donde $\varepsilon_i \sim N(0, \sigma^2)$ y son independientes; $i = 1, 2, \dots, n$, $n > 2$ (ver Apéndice B). Con espacio paramétrico $\Omega = \{(\alpha, \beta, \sigma^2) | \alpha \in \mathfrak{R}, \beta \in \mathfrak{R}, \sigma^2 > 0\}$. Esto se puede escribir como

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}, \quad \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}.$$

Luego

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix}, \quad \mathbf{X}'\mathbf{Y} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix}$$

$$(\mathbf{X}'\mathbf{X})^{-1} = \left(\frac{1}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \right) \begin{pmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{pmatrix}.$$

Por lo tanto

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \left(\frac{1}{n \sum_{i=1}^n (x_i - \bar{x})^2} \right) \begin{bmatrix} \left(\sum_{i=1}^n y_i \right) \left(\sum_{i=1}^n x_i^2 \right) - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n x_i y_i \right) \\ n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right) \end{bmatrix}.$$

Observándose que si

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad y \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i.$$

Entonces se tiene que

$$\begin{aligned} n \sum_{i=1}^n (x_i - \bar{x})^2 &= n \sum_{i=1}^n (x_i - 2x_i\bar{x} + \bar{x}^2) = n \left(\sum_{i=1}^n x_i - 2\bar{x} \sum_{i=1}^n x_i + n\bar{x}^2 \right) \\ &= n \left(\sum_{i=1}^n x_i - \frac{2}{n} \left(\sum_{i=1}^n x_i \right)^2 + n \left(\sum_{i=1}^n x_i \right)^2 \right) = n \sum_{i=1}^n x_i - \left(\sum_{i=1}^n x_i \right)^2. \end{aligned}$$

Similarmente, notando que

$$\begin{aligned} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) &= \sum_{i=1}^n (x_i y_i - x_i \bar{y} - \bar{x} y_i + \bar{x} \bar{y}) \\ &= \sum_{i=1}^n x_i y_i - \bar{y} \sum_{i=1}^n x_i - \bar{x} \sum_{i=1}^n y_i + n\bar{x}\bar{y} \\ &= \sum_{i=1}^n x_i y_i - \frac{1}{n} \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right) - \frac{1}{n} \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right) \\ &\quad + \frac{1}{n} \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right) \\ &= \sum_{i=1}^n x_i y_i - \frac{1}{n} \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right). \end{aligned}$$

De donde

$$n \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right).$$

Por lo tanto

$$\hat{\beta} = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} \bar{y} - \hat{\beta}\bar{x} \\ \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \end{pmatrix}.$$

También, para $\hat{\sigma}^2$ se tiene

$$\begin{aligned}
 (n-2)\hat{\sigma}^2 &= (\mathbf{y}\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}) \\
 &= \sum_{i=1}^n y_i - \left(\bar{y} - \hat{\beta}\bar{x}, \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{pmatrix} \\
 &= \sum_{i=1}^n y_i - (\bar{y} - \hat{\beta}\bar{x}) \sum_{i=1}^n y_i - \hat{\beta} \sum_{i=1}^n x_i y_i \\
 &= \sum_{i=1}^n (y_i - \bar{y})^2 - \frac{(\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}))^2}{\sum_{i=1}^n (x_i - \bar{x})^2}.
 \end{aligned}$$

Aplicando estos resultados a los R modelos lineales simples se obtiene que

$$\hat{\boldsymbol{\beta}}_{\rho} = \begin{pmatrix} \hat{\alpha}_{\rho} \\ \hat{\beta}_{\rho} \end{pmatrix} = \begin{pmatrix} \bar{y}_{\rho} - \hat{\beta}_{\rho}\bar{x}_{\rho} \\ \frac{\sum_{i=1}^n (x_{i\rho} - \bar{x}_{\rho})(y_{i\rho} - \bar{y}_{\rho})}{\sum_{i=1}^n (x_{i\rho} - \bar{x}_{\rho})^2} \end{pmatrix}, \quad \rho = 1, 2, \dots, R$$

$$\hat{\sigma}^2 = \frac{\sum_{\rho=1}^R (n_{\rho} - 2)\hat{\sigma}_{\rho}^2}{\sum_{\rho=1}^R (n_{\rho} - 2)}$$

Donde

$$\hat{\sigma}_{\rho}^2 = \frac{1}{n_{\rho} - 2} \left\{ \sum_{j=1}^{n_{\rho}} (y_{\rho j} - \bar{y}_{\rho})^2 - \frac{[\sum_{j=1}^{n_{\rho}} (y_{\rho j} - \bar{y}_{\rho})(x_{\rho j} - \bar{x}_{\rho})]^2}{\sum_{j=1}^{n_{\rho}} (x_{\rho j} - \bar{x}_{\rho})} \right\}$$

Por lo tanto, los estadísticos de prueba de las hipótesis *i.- iii.* serán una función de estos $R + 1$ estimadores completos y suficientes. Con tal objetivo en mente, considere la distribución de la variable aleatoria Z_{ρ} , donde $Z_{\rho} = a\hat{\alpha} + b\hat{\beta}$

donde a y b son constantes conocidas a ser especificadas posteriormente.

Recuérdese que $\widehat{\boldsymbol{\beta}}_\rho \sim N_2(\boldsymbol{\beta}_\rho, \hat{\sigma}^2_\rho(\mathbf{X}'\mathbf{X})^{-1})$ y nótese que $Z_\rho = (a, b)\widehat{\boldsymbol{\beta}}_\rho$, por lo tanto,

$$Z_\rho \sim N\left((a, b)\boldsymbol{\beta}_\rho, \sigma^2_\rho(a, b)(\mathbf{X}'\mathbf{X})^{-1}\begin{pmatrix} a \\ b \end{pmatrix}\right).$$

Más aún,

$$E(Z_\rho) = E\left((a, b)\widehat{\boldsymbol{\beta}}_\rho\right) = (a, b)\begin{pmatrix} \alpha_\rho \\ \beta_\rho \end{pmatrix} = a\alpha_\rho + b\beta_\rho, y$$

$$\begin{aligned} \text{Var}(Z_\rho) &= \text{Var}\left((a, b)\widehat{\boldsymbol{\beta}}_\rho\right) = (a, b)\text{Var}\widehat{\boldsymbol{\beta}}_\rho\begin{pmatrix} a \\ b \end{pmatrix} = (a, b)\sigma^2(\mathbf{X}'\mathbf{X})^{-1}\begin{pmatrix} a \\ b \end{pmatrix} \\ &= \sigma^2(a, b)\begin{pmatrix} \sum_{i=1}^n x_{i\rho}^2 & -\sum_{i=1}^n x_{i\rho} \\ -\sum_{i=1}^n x_{i\rho} & n_\rho \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}\left(\frac{1}{n\sum_{i=1}^n (x_{i\rho} - \bar{x}_\rho)^2}\right) \\ &= \frac{\sigma^2\left(a^2\sum_{i=1}^n x_{i\rho}^2 - 2ab\sum_{i=1}^n x_{i\rho} + b^2n_\rho\right)}{n\sum_{i=1}^n (x_{i\rho} - \bar{x}_\rho)^2} = \frac{\sum_{i=1}^n (ax_{i\rho} - b)^2}{n\sum_{i=1}^n (x_{i\rho} - \bar{x}_\rho)^2}. \end{aligned}$$

Ahora considérese la variable aleatoria W definida como

$$W = \frac{\sum_{\rho=1}^R (Z_\rho - \bar{Z}^*)^2 d_{\rho\rho}}{(R-1)\hat{\sigma}^2}.$$

Donde

$$\bar{Z}^* = \frac{\sum_{\rho=1}^R Z_\rho d_{\rho\rho}}{\sum_{\rho=1}^R d_{\rho\rho}}, \quad d_{\rho\rho} = \frac{n_\rho \sum_{i=1}^{n_\rho} (x_{i\rho} - \bar{x}_\rho)^2}{\sum_{i=1}^{n_\rho} (ax_{j\rho} - b)^2}.$$

Nótese que

$$u_1 = \frac{\sum_{\rho=1}^R (Z_\rho - \bar{Z}^*)^2 d_{\rho\rho}}{\hat{\sigma}^2} \sim \chi_{R-1, \lambda}^2.$$

Donde

$$\lambda = \left(\frac{1}{2\sigma^2} \right) \sum_{\rho=1}^R (\mu_\rho - \bar{\mu}^*)^2 d_{\rho\rho}, \quad \bar{\mu}^* = \frac{\sum_{\rho=1}^R \mu_\rho d_{\rho\rho}}{\sum_{\rho=1}^R d_{\rho\rho}}.$$

Definiendo $u_2 = (N - 2R) \hat{\sigma}^2 / \sigma^2$,

$$u_2 \sim \chi_{N-2R}^2.$$

Más aún, u_1 y u_2 son independientes, por lo tanto

$$W \sim F_{R-1, N-2R: \lambda}.$$

Y obsérvese que $\lambda = 0$, sí y solo sí, $\mu_1 = \mu_2 = \dots = \mu_R$. Así, sustituyendo μ_ρ por $a\alpha_\rho + b\beta_\rho$, se tiene el siguiente resultado:

TEOREMA 1.1. *Considérense los R modelos lineales simples definidos en la ecuación (1). Entonces el estadístico W dado por*

$$W = \frac{\sum_{\rho=1}^R \left[a\hat{\alpha}_\rho + b\hat{\beta}_\rho - \frac{\sum_{i=1}^R (a\hat{\alpha}_i + b\hat{\beta}_i) d_{ii}}{\sum_{j=1}^R d_{jj}} \right]^2 d_{\rho\rho}}{(R-1)\hat{\sigma}^2}.$$

Es distribuido $F(W: R-1, N-2R: \lambda)$ con

$$\lambda = \left(\frac{1}{2\sigma^2} \right) \sum_{\rho=1}^R \left[a\alpha_\rho + b\beta_\rho - \frac{\sum_{i=1}^R (a\alpha_i + b\beta_i) d_{ii}}{\sum_{j=1}^R d_{jj}} \right]^2 d_{\rho\rho}.$$

Donde

$$d_{\rho\rho} = \frac{n_{\rho} \sum_{j=1}^{n_{\rho}} (x_{j\rho} - \bar{x}_{\rho})^2}{\sum_{i=1}^{n_{\rho}} (ax_{i\rho} - b)^2}.$$

Nótese que $\lambda = 0$ sí y solo sí $a\alpha_1 + b\beta_1 = a\alpha_2 + b\beta_2 = \dots = a\alpha_R + b\beta_R$.

Como consecuencia, se tiene el siguiente resultado:

COROLARIO 1. *Considérense los R modelos lineales simples definidos en la ecuación (1). La prueba de razón de verosimilitud de tamaño α para la hipótesis H_0 versus H_a es la siguiente: se rechaza H_0 sí y solo sí*

$$W \geq F_{\alpha; R-1, N-2R}$$

Donde W es definido en el teorema 1 y H_0 y H_a son:

$$H_0: a\alpha_1 + b\beta_1 = a\alpha_2 + b\beta_2 = \dots = a\alpha_R + b\beta_R$$

H_a : al menos una igualdad es una desigualdad

Prueba: es inmediata a partir del teorema 1.

A continuación, se demuestra como este corolario es usado para probar las hipótesis *i.- iii.*

TEOREMA 1.2. *Considérense los R modelos lineales simples dados en la ecuación (1). Las pruebas de hipótesis *i.- iii.* están dadas a continuación:*

- i.* $H_0: \beta_1 = \beta_2 = \dots = \beta_R$ (las líneas son paralelas) versus $H_a: \beta_i \neq \beta_j$ para al menos una $i \neq j$ (no todas las R líneas son paralelas). La regla de decisión es: se rechaza H_0 sí y solo sí $W_{\rho} \geq F_{\alpha; R-1, N-2R}$ (ver Apéndice C), donde

$$W_\rho = \frac{\sum_{\rho=1}^R \left[\hat{\beta}_\rho - \frac{\sum_{j=1}^R \hat{\beta}_j b_{jj}}{\sum_{i=1}^R b_{ii}} \right]^2 b_{\rho\rho}}{(R-1)\hat{\sigma}^2},$$

donde,

$$b_{\rho\rho} = \sum_{i=1}^{n_\rho} (x_{i\rho} - \bar{x}_\rho)^2.$$

- ii. $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_R$ (las R líneas tienen el mismo intercepto) versus $H_a: \alpha_i \neq \alpha_j$ para al menos una $i \neq j$.

$$W_I = \frac{\sum_{\rho=1}^R \left[\hat{\alpha}_\rho - \frac{\sum_{j=1}^R \hat{\alpha}_j a_{jj}}{\sum_{i=1}^R a_{ii}} \right]^2 a_{\rho\rho}}{(R-1)\hat{\sigma}^2},$$

donde,

$$a_{\rho\rho} = \frac{n_\rho \sum_{i=1}^{n_\rho} (x_{i\rho} - \bar{x}_\rho)^2}{\sum_{j=1}^{n_\rho} x_{j\rho}^2}.$$

La prueba de tamaño α de H_0 versus H_a es: se rechaza H_0 sí y solo sí $W_I \geq$

$F_{\alpha; R-1, N-2R}$.

- iii. $H_0: \alpha_1 + \beta_1 x_0 = \alpha_2 + \beta_2 x_0 = \dots = \alpha_R + \beta_R x_0$ (las R líneas se intersecan en $x = x_0$ conocida) versus H_a : al menos una igualdad es una desigualdad (las R líneas no se intersecan en $x = x_0$). La prueba de tamaño α para H_0 versus H_a es: se rechaza H_0 sí y solo sí $W_o \geq F_{\alpha; R-1, N-2R}$, donde

$$W_o = \frac{\sum_{\rho=1}^R \left[\hat{\alpha}_\rho + \hat{\beta}_\rho x_0 - \frac{\sum_{i=1}^R (\hat{\alpha}_i + \hat{\beta}_i x_0) c_{ii}}{\sum_{j=1}^R c_{jj}} \right]^2 c_{\rho\rho}}{(R-1)\hat{\sigma}^2},$$

donde

$$c_{\rho\rho} = \frac{n_{\rho} \sum_{i=1}^{n_{\rho}} (x_{i\rho} - \bar{x}_{\rho})^2}{\sum_{j=1}^{n_{\rho}} (x_{j\rho} - \bar{x}_0)^2}.$$

Prueba: esta es inmediata a partir del corolario 1 para probar si las R líneas son paralelas. Defina $a = 0$ y $b = 1$, para probar que las R líneas se intersecan en $x = x_0$, defina $a = 0$ y $b = x_0$; finalmente, para probar que las R líneas tienen el mismo intercepto, tome $a = 0$ y $b = 0$, probando así los tres resultados deseados.

Prueba para la Igualdad de un Conjunto de Modelos Lineales

En esta sección es derivado el estadístico de prueba para probar la hipótesis de que R modelos lineales generales son iguales.

TEOREMA 1.3. Sean $y_{\rho} = X_{\rho}\beta_{\rho} + \varepsilon_{\rho}, \rho = 1, 2, \dots, R, R$ modelos lineales generales tales que $X_{\rho} \in \mathfrak{R}^{n_{\rho} \times p}, \varepsilon_{\rho} \in \mathfrak{R}^{n_{\rho}}$ son vectores aleatorios tales que $\varepsilon_{\rho} \sim N_{n_{\rho}}(\mathbf{0}, \sigma^2 \mathbf{I})$. Una prueba de tamaño α para la hipótesis

$$H_0: \beta_1 = \beta_2 = \dots = \beta_R$$

versus

H_a : al menos una igualdad no se cumple.

Esta dada por:

Se rechaza H_0 sí y sólo sí $W \geq F_{\alpha, (R-1)_p, N-R_p}$ (ver Apéndice C), donde

$$N = \sum_{\rho=1}^R n_{\rho},$$

y

$$W = \left(\frac{\sum_{\rho=1}^R \mathbf{y}_{\rho}' (\mathbf{X}_{\rho}' \mathbf{X}_{\rho}^{-}) \mathbf{y}_{\rho} - (\sum_{i=1}^R \mathbf{y}_i' \mathbf{X}_i) (\sum_{\rho=1}^R \mathbf{X}_{\rho}' \mathbf{X}_{\rho}^{-})^{-1} (\sum_{j=1}^R \mathbf{X}_j' \mathbf{y}_j)}{\sum_{\rho=1}^R \mathbf{y}_{\rho}' \mathbf{y}_{\rho} - \sum_{\rho=1}^R \mathbf{y}_{\rho}' (\mathbf{X}_{\rho} \mathbf{X}_{\rho}^{-}) \mathbf{y}_{\rho}} \right) \left(\frac{N - R\rho}{(R - 1)\rho} \right).$$

Prueba. Considere las siguientes matrices y vectores particionados

$$\mathbb{Y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_R \end{bmatrix}, \quad \mathbb{X} = \begin{bmatrix} \mathbf{X}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{X}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{X}_R \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_R \end{bmatrix}, \quad \mathbb{E} = \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_R \end{bmatrix}.$$

Dado que los vectores $\boldsymbol{\varepsilon}_{\rho}$ se distribuyen independientemente normales, se tiene que $\mathbb{E} \sim N_N(\mathbf{0}, \sigma^2 \mathbf{I})$. Además $\mathbb{X} \in \mathfrak{R}^{N \times R\rho}$, $\mathbb{Y} \in \mathfrak{R}^N$ y $\boldsymbol{\beta} \in \mathfrak{R}^{R\rho}$. Se puede escribir el modelo lineal de la siguiente manera:

$$\mathbb{Y} = \mathbb{X}\boldsymbol{\beta} + \mathbb{E}, \quad \mathbb{E} \sim N_N(\mathbf{0}, \sigma^2 \mathbf{I}). \quad (1)$$

El cual es equivalente a los N modelos lineales

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{X}_1 \boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_1 \\ \mathbf{y}_2 &= \mathbf{X}_2 \boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}_2 \\ &\vdots \\ \mathbf{y}_R &= \mathbf{X}_R \boldsymbol{\beta}_R + \boldsymbol{\varepsilon}_R. \end{aligned}$$

Por lo tanto, se desea probar la hipótesis $H_0: \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_R$ en el modelo lineal general definido en la ecuación (1). Por lo tanto, se puede aplicar la razón de verosimilitud generalizada, con tal propósito es necesario determinar $\tilde{\sigma}_{\omega}^2$ y $\tilde{\sigma}_{\Omega}^2$ para el modelo (1), luego

$$W = \left(\frac{N - R\rho}{(R - 1)\rho} \right) \left(\frac{\tilde{\sigma}_{\omega}^2 - \tilde{\sigma}_{\Omega}^2}{\tilde{\sigma}_{\Omega}^2} \right)$$

Bajo el modelo completo se tiene que $\mathbb{X}'\mathbb{X}\hat{\boldsymbol{\beta}} = \mathbb{X}'\mathbb{Y}$, por lo tanto $\hat{\boldsymbol{\beta}} = (\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}'\mathbb{Y}$. Luego, nótese que

$$\begin{aligned}\mathbb{X}'\mathbb{X} &= \begin{bmatrix} \mathbf{X}'_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}'_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}'_R \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_R \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{X}'_1\mathbf{X}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}'_2\mathbf{X}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}'_R\mathbf{X}_R \end{bmatrix}.\end{aligned}$$

Luego

$$(\mathbb{X}'\mathbb{X})^{-1} = \begin{bmatrix} (\mathbf{X}'_1\mathbf{X}_1)^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & (\mathbf{X}'_2\mathbf{X}_2)^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & (\mathbf{X}'_R\mathbf{X}_R)^{-1} \end{bmatrix}.$$

También

$$\mathbb{X}'\mathbb{Y} = \begin{bmatrix} \mathbf{X}'_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}'_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}'_R \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_R \end{bmatrix} = \begin{bmatrix} \mathbf{X}'_1\mathbf{y}_1 \\ \mathbf{X}'_2\mathbf{y}_2 \\ \vdots \\ \mathbf{X}'_R\mathbf{y}_R \end{bmatrix}.$$

De donde

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_R \end{bmatrix} \\ &= \begin{bmatrix} (\mathbf{X}'_1\mathbf{X}_1)^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & (\mathbf{X}'_2\mathbf{X}_2)^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & (\mathbf{X}'_R\mathbf{X}_R)^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{X}'_1\mathbf{y}_1 \\ \mathbf{X}'_2\mathbf{y}_2 \\ \vdots \\ \mathbf{X}'_R\mathbf{y}_R \end{bmatrix}\end{aligned}$$

$$\begin{aligned}
&= \begin{bmatrix} (\mathbf{X}_1' \mathbf{X}_1)^{-1} \mathbf{X}_1' \mathbf{y}_1 \\ (\mathbf{X}_2' \mathbf{X}_2)^{-1} \mathbf{X}_2' \mathbf{y}_2 \\ \vdots \\ (\mathbf{X}_R' \mathbf{X}_R)^{-1} \mathbf{X}_R' \mathbf{y}_R \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{X}_1^- \mathbf{y}_1 \\ \mathbf{X}_2^- \mathbf{y}_2 \\ \vdots \\ \mathbf{X}_R^- \mathbf{y}_R \end{bmatrix}.
\end{aligned}$$

Entonces, obsérvese que $\hat{\boldsymbol{\beta}}_\rho$ es el mismo obtenido desde el ρ -ésimo modelo lineal general $\mathbf{y}_\rho = \mathbf{X}_\rho \boldsymbol{\beta}_\rho + \boldsymbol{\varepsilon}_\rho, \rho = 1, 2, \dots, R$. Además, obsérvese que

$$\begin{aligned}
\hat{\boldsymbol{\beta}} \mathbb{X}' \mathbb{Y} &= (\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \dots, \hat{\boldsymbol{\beta}}_R) \begin{bmatrix} \mathbf{X}_1' \mathbf{y}_1 \\ \mathbf{X}_2' \mathbf{y}_2 \\ \vdots \\ \mathbf{X}_R' \mathbf{y}_R \end{bmatrix} \\
&= \sum_{\rho=1}^R \hat{\boldsymbol{\beta}}_\rho' \mathbf{X}_\rho' \mathbf{y}_\rho = \sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho \mathbf{X}_\rho^- \mathbf{y}_\rho.
\end{aligned}$$

Así, se tiene la siguiente expresión para $\tilde{\sigma}_\Omega^2$

$$\begin{aligned}
N \tilde{\sigma}_\Omega^2 &= \|\hat{\mathbb{E}}\|^2 = \|\mathbb{Y} - \mathbb{X} \hat{\boldsymbol{\beta}}\|^2 = \mathbb{Y}' \mathbb{Y} - \hat{\boldsymbol{\beta}}' \mathbb{X}' \mathbb{Y} = \sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{y}_\rho - \sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho \mathbf{X}_\rho^- \mathbf{y}_\rho \\
&= \sum_{\rho=1}^R \mathbf{y}_\rho' (\mathbf{I}_{n_\rho} - \mathbf{X}_\rho \mathbf{X}_\rho^-) \mathbf{y}_\rho.
\end{aligned}$$

Ahora, el modelo reducido por la hipótesis $H_0: \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \dots = \boldsymbol{\beta}_R = \boldsymbol{\delta} \in \mathfrak{R}^\rho$ y se escribe como

$$\mathbb{Y} = \mathbb{X} \boldsymbol{\nu} + \mathbb{E}.$$

Donde

$$\mathbf{v} = \begin{bmatrix} \boldsymbol{\delta}^1 \\ \boldsymbol{\delta}^2 \\ \vdots \\ \boldsymbol{\delta}^R \end{bmatrix} \in \mathfrak{R}^{R\rho}.$$

Así se tiene que

$$\begin{aligned} N\tilde{\sigma}_\omega^2 &= \min(\|\mathbb{E}\|^2) = \min(\mathbb{E}'\mathbb{E}) = \min\left(\sum_{\rho=1}^R \boldsymbol{\varepsilon}_\rho' \boldsymbol{\varepsilon}_\rho\right) \\ &= \min\left[\sum_{\rho=1}^R (\mathbf{y}_\rho - \mathbf{X}_\rho \boldsymbol{\delta})' (\mathbf{y}_\rho - \mathbf{X}_\rho \boldsymbol{\delta})\right] \\ &= \min\left[\sum_{\rho=1}^R (\mathbf{y}_\rho' \mathbf{y}_\rho - \mathbf{y}_\rho' \mathbf{X}_\rho \boldsymbol{\delta} - \boldsymbol{\delta}' \mathbf{X}_\rho' \mathbf{y}_\rho + \boldsymbol{\delta}' \mathbf{X}_\rho' \mathbf{X}_\rho \boldsymbol{\delta})\right] \\ &= \min\left[\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{y}_\rho - 2 \sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho \boldsymbol{\delta} + \boldsymbol{\delta}' \left(\sum_{\rho=1}^R (\mathbf{X}_\rho' \mathbf{X}_\rho) \boldsymbol{\delta}\right)\right]. \end{aligned} \quad (2)$$

Diferenciando, se tiene que

$$\begin{aligned} dN\tilde{\sigma}_\omega^2 &= -2\left(\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho\right) d\boldsymbol{\delta} + d\boldsymbol{\delta}' \left(\sum_{\rho=1}^R (\mathbf{X}_\rho' \mathbf{X}_\rho) \boldsymbol{\delta}\right) + \left(\boldsymbol{\delta}' \sum_{\rho=1}^R (\mathbf{X}_\rho' \mathbf{X}_\rho)\right) d\boldsymbol{\delta} \\ &= -2\left(\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho\right) d\boldsymbol{\delta} + 2\left(\boldsymbol{\delta}' \sum_{\rho=1}^R (\mathbf{X}_\rho' \mathbf{X}_\rho)\right) d\boldsymbol{\delta}. \end{aligned} \quad (3)$$

Por lo tanto

$$\frac{N\partial\tilde{\sigma}_\omega^2}{\partial\boldsymbol{\delta}} = -2\left(\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho\right)' + 2\left(\boldsymbol{\delta}' \sum_{\rho=1}^R (\mathbf{X}_\rho' \mathbf{X}_\rho)\right)'.$$

Igualando a cero y denotando la solución como $\hat{\boldsymbol{\delta}}$

$$-2 \sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{y}_\rho + 2 \left(\sum_{\rho=1}^R (\mathbf{X}_\rho' \mathbf{X}_\rho) \right) \hat{\boldsymbol{\delta}} = 0.$$

Por lo tanto,

$$\hat{\boldsymbol{\delta}} = \left(\sum_{\rho=1}^R (\mathbf{X}_\rho' \mathbf{X}_\rho) \right)^{-1} \left(\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{y}_\rho \right) \quad (4)$$

Veamos que $N\tilde{\sigma}_\omega^2$ tiene un mínimo en la ecuación (4). Para ello, de la ecuación (3) se tiene que

$$d^2 N\tilde{\sigma}_\omega^2 = 2d\boldsymbol{\delta}' \left(\sum_{\rho=1}^R (\mathbf{X}_\rho' \mathbf{X}_\rho) \right) d\boldsymbol{\delta}.$$

De donde la matriz hessiana está dada por

$$\frac{N\partial^2 \tilde{\sigma}_\omega^2}{\partial \boldsymbol{\delta} \partial \boldsymbol{\delta}'} = 2 \sum_{\rho=1}^R (\mathbf{X}_\rho' \mathbf{X}_\rho).$$

Luego, nótese que $r(\mathbf{X}_\rho) = r(\mathbf{X}_\rho' \mathbf{X}_\rho) = \rho$ además, $\mathbf{X}_\rho' \mathbf{X}_\rho \in \mathfrak{R}^{\rho \times \rho}$.

Es decir $\mathbf{X}_\rho' \mathbf{X}_\rho > \mathbf{0}$ para todo $\rho = 1, 2, \dots, R$ y, por lo tanto,

$$\frac{N\partial^2 \tilde{\sigma}_\omega^2}{\partial \boldsymbol{\delta} \partial \boldsymbol{\delta}'} = 2 \sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{X}_\rho > \mathbf{0}.$$

De donde se concluye que $N\tilde{\sigma}_\omega^2$ tiene su mínimo en $\hat{\boldsymbol{\delta}}$, dado en la ecuación (4). Sustituyendo este último en $N\tilde{\sigma}_\omega^2$, ecuación (2), se obtiene que

$$\begin{aligned}
N\tilde{\sigma}_\omega^2 &= \sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{y}_\rho - 2 \left(\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho \right) \left(\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{X}_\rho \right)^{-1} \left(\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{y}_\rho \right) \\
&\quad + \left(\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho \right) \left(\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{X}_\rho \right)^{-1} \left(\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{X}_\rho \right) \left(\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{X}_\rho \right)^{-1} \left(\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{y}_\rho \right), \\
&= \sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{y}_\rho - \left(\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho \right) \left(\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{X}_\rho \right)^{-1} \left(\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{y}_\rho \right).
\end{aligned}$$

De donde se obtiene que

$$\begin{aligned}
N\tilde{\sigma}_\omega^2 &= \sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{y}_\rho - \left(\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho \right) \left(\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{X}_\rho \right)^{-1} \left(\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{y}_\rho \right) - \sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho \\
&\quad + \sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho \mathbf{X}_\rho^{-1} \mathbf{y}_\rho - \left(\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho \right) \left(\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{X}_\rho \right)^{-1} \left(\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{y}_\rho \right).
\end{aligned}$$

Por lo tanto

$$\begin{aligned}
W &= \left(\frac{N - R\rho}{(R - 1)\rho} \right) \left(\frac{\tilde{\sigma}_\omega^2 - \tilde{\sigma}_\Omega^2}{\tilde{\sigma}_\Omega^2} \right) \\
&= \left(\frac{N - R\rho}{(R - 1)\rho} \right) \left(\frac{\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho \mathbf{X}_\rho^{-1} \mathbf{y}_\rho - (\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho) (\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{X}_\rho)^{-1} (\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{y}_\rho)}{\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{y}_\rho - \sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho \mathbf{X}_\rho^{-1} \mathbf{y}_\rho} \right).
\end{aligned}$$

Finalmente, obsérvese que

$$W \sim F_{(R-1)\rho, N-R\rho, \lambda},$$

donde

$$\lambda = \left(\frac{1}{2\sigma^2} \right) \left[\sum_{\rho=1}^R \beta_{\rho}' X_{\rho} X_{\rho}' \beta_{\rho} - \left(\sum_{\rho=1}^R \beta_{\rho}' X_{\rho}' X_{\rho} \right) \left(\sum_{\rho=1}^R X_{\rho}' X_{\rho} \right)^{-1} \left(\sum_{\rho=1}^R X_{\rho}' X_{\rho} \beta_{\rho} \right) \right].$$

Con lo cual el teorema está demostrado.

Aplicación

La prueba de igualdad de modelos es aplicada en el siguiente experimento agrícola donde se estimaron las curvas de crecimiento poblacional de tres plagas del suelo en el cultivo del maíz en el Bajío de Guanajuato, con el propósito de eficientar su manejo integrado en esta región.

La intención del estudio es aportar información sobre las densidades poblacionales de *D. undecimpunctata*, *D. balteata* y *D. virgifera*, que pueda ser considerada para generar mejores sistemas de manejo y control integrado de estas especies en el cultivo de maíz en Irapuato, Guanajuato, México.

Los objetivos planteados en la investigación son:

- (1) Proponer para las tres especies de *Diabrotica*, formas funcionales iguales para los modelos de curvas de crecimiento poblacional;
- (2) Estimar estas curvas de crecimiento poblacional mediante técnicas de regresión;
- (3) Comparar las curvas de crecimiento que modelan la densidad poblacional de los adultos de las tres especies bajo estudio en función de las variables climáticas de temperatura y precipitación con registros del promedio de diez años; y

- (4) Estimar los valores de la temperatura y la precipitación para los cuales se presenta la máxima población de individuos adultos para cada una de las tres especies de coleópteros.

Se calculó el promedio mensual de los individuos adultos de las tres especies y se tomaron como variables dependientes (Y), y como variables independientes la temperatura mensual promedio (x_1) y el promedio mensual de precipitación pluvial acumulada (x_2) de 1993 a 2002.

Se estimó el mejor modelo de regresión para cada una de las especies (variable dependiente Y). En este sentido, se consideraron los valores de la prueba de F de los análisis de la varianza de las regresiones y los coeficientes de determinación parciales (R^2) como criterios de selección, ésto elaborado en el paquete SAS versión 9.0. Una vez determinado este proceso, se obtuvo el siguiente polinomio de tercer grado:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1^2 + \beta_4x_2^2 + \beta_5x_1x_2 + \beta_6x_1x_2^2$$

donde

x_1 = Temperatura promedio mensual.

x_2 = Precipitación pluvial promedio mensual.

Se procedió a estimar los tres polinomios para las curvas de crecimiento poblacional para la cantidad de adultos de las tres especies de *Diabrotica* (Y) en función de los promedios mensuales de temperatura (x_1) y precipitación pluvial acumulada (x_2) durante diez años.

$$y_1 = -102.8966 + 17.5947x_1 - 7.4358x_2 - 0.6973x_1^2 + 0.0886x_2^2 + 0.4687x_1x_2 - 0.0049x_1x_2^2$$

$$y_2 = -111.9586 + 15.5984x_1 - 4.6132x_2 - 0.5420x_1^2 + 0.0715x_2^2 + 0.3136x_1x_2 - 0.0040x_1x_2^2$$

$$y_3 = -3.5830 + 0.5435x_1 - 0.1434x_2 - 0.0197x_1^2 + 0.0016x_2^2 + 0.0090x_1x_2 - 0.000091x_1x_2^2$$

donde

y_1 = Valor numérico estimado de adultos de la especie *Diabrotica undecimpunctata*.

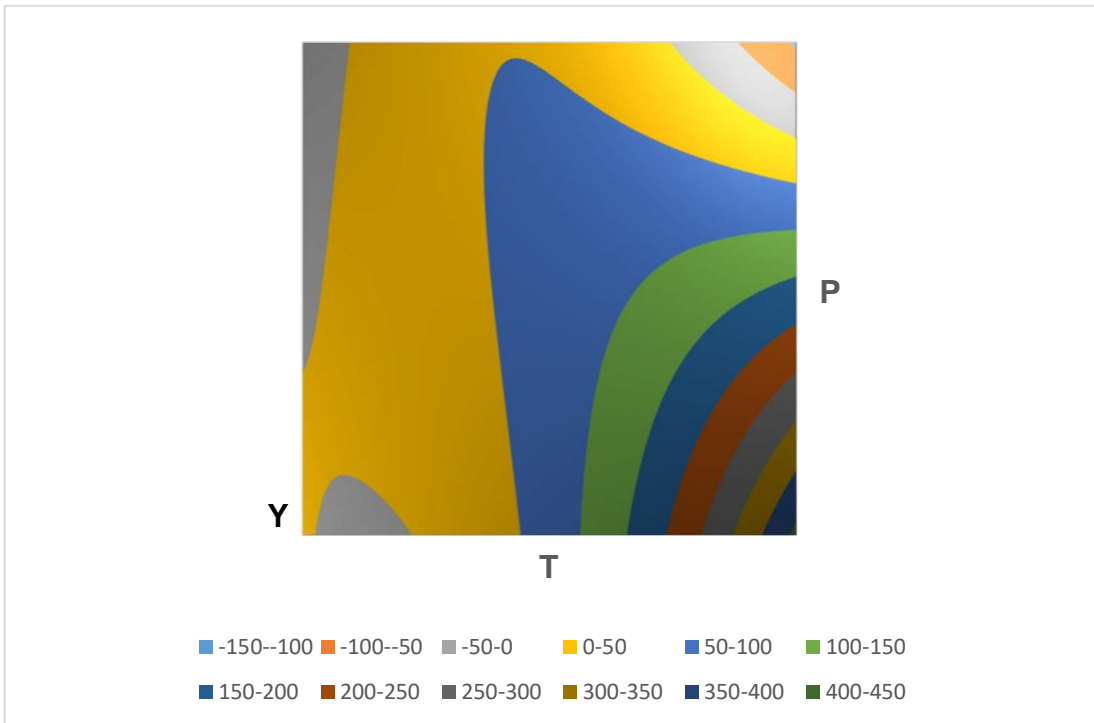
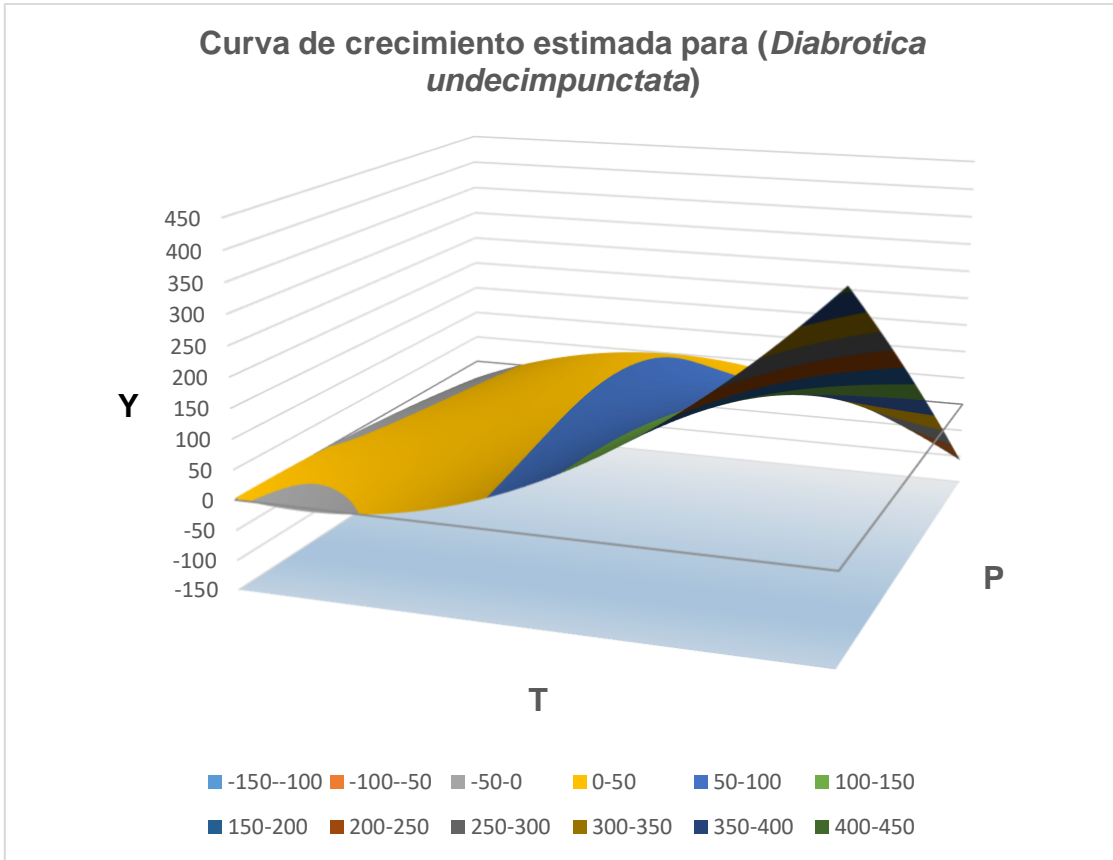
y_2 = Valor numérico estimado de adultos de la especie *Diabrotica balteata*.

y_3 = Valor numérico estimado de adultos de la especie *Diabrotica virgifera*.

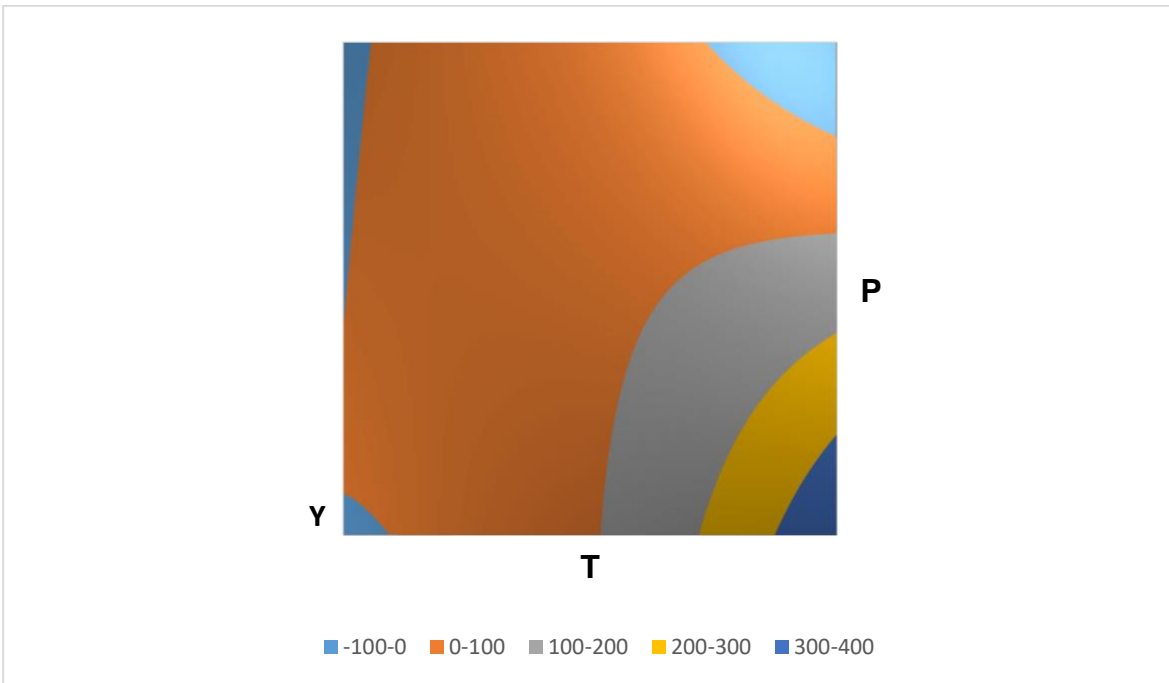
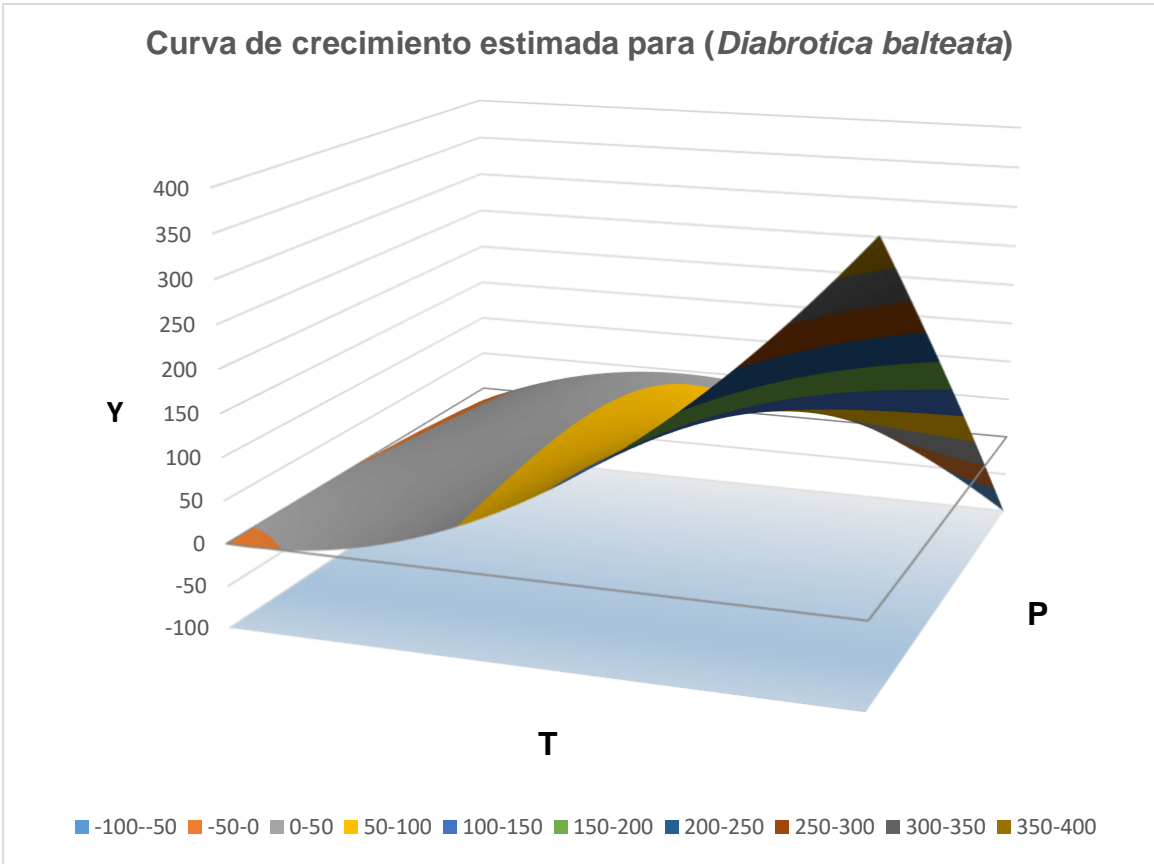
Luego, se realizó la comparación entre sí de las tres curvas de crecimiento poblacionales con forma funcional de polinomios de tercer grado, es decir, la misma forma funcional, las cuales se muestran en las Figuras 3, 4 y 5.

En estas gráficas de curvas de crecimiento estimadas se pueden observar los números máximos de adultos de cada especie, para cada correspondiente punto crítico de temperatura y precipitación.

Y finalmente, mediante técnicas estándar de programación matemática aplicando un método cuasi-Newton, se determinaron los puntos críticos de temperatura y precipitación (x_1, x_2) para cada curva de crecimiento, que permitieron estimar el número máximo de individuos de cada especie a través de la subrutina nlminb del software estadístico R, la cual permite optimizar funciones de varias variables sin restricciones o con restricciones tipo caja, (ver Apéndice D).

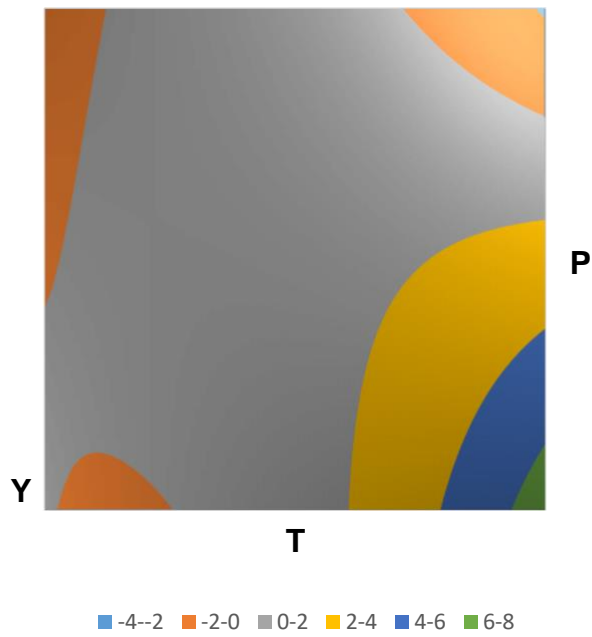
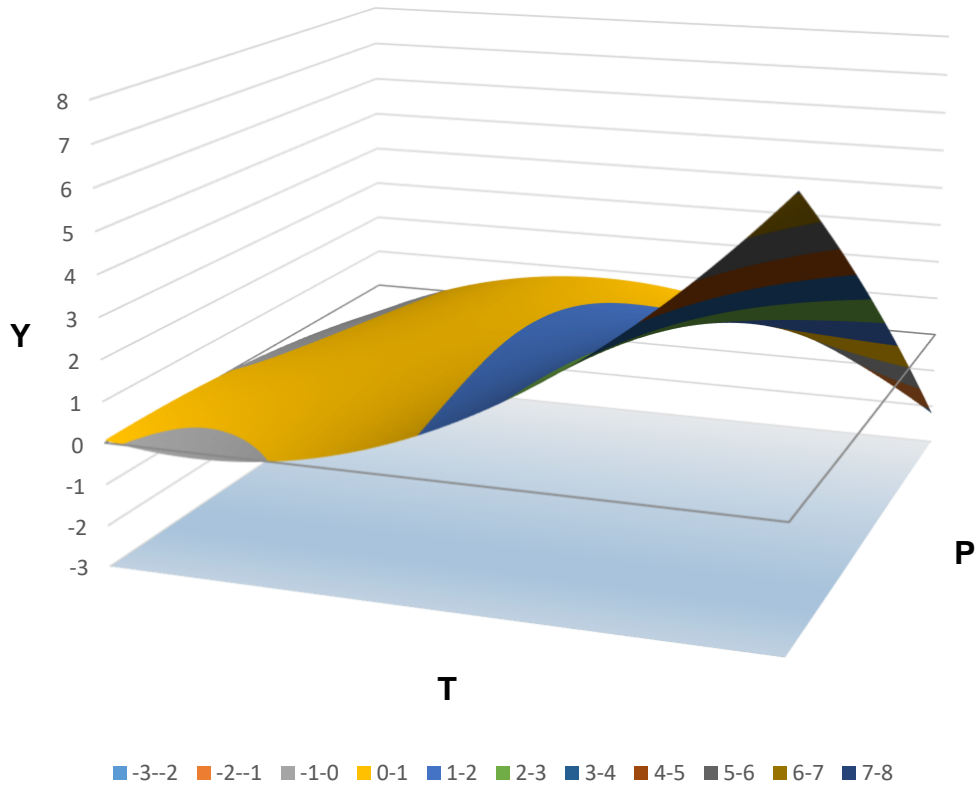


Gráfica 1. Curva de crecimiento estimada para *Diabrotica undecimpunctata*.



Gráfica 2. Curva de crecimiento estimada para *Diabrotica balteata*.

Curva de crecimiento estimada para (*Diabrotica virgifera*)



Gráfica 3. Curva de crecimiento estimada para *Diabrotica virgifera*.

En los Cuadros 1, 2 y 3 se presentan los resultados obtenidos de los análisis de varianza para cada una de las poblaciones de *Diabrotica* estudiadas. El Cuadro 4 muestra un resumen de los resultados obtenidos.

Para la comparación de los modelos entre sí, se utilizó la prueba de hipótesis, de acuerdo con lo visto en el Teorema 1.3 de este capítulo,

$$H_0: \beta_1 = \beta_2 = \dots = \beta_R$$

versus

H_a : al menos una igualdad no se cumple.

Está dada por:

Se rechaza H_0 si y sólo si $W \geq F_{\alpha, (R-1)\rho, N-R\rho}$.

El valor del estadístico de prueba se obtuvo de la siguiente manera:

$$\begin{aligned} W &= \left(\frac{N - R\rho}{(R - 1)\rho} \right) \left(\frac{\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho \mathbf{X}_\rho^{-1} \mathbf{y}_\rho - (\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho) (\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{X}_\rho)^{-1} (\sum_{\rho=1}^R \mathbf{X}_\rho' \mathbf{y}_\rho)}{\sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{y}_\rho - \sum_{\rho=1}^R \mathbf{y}_\rho' \mathbf{X}_\rho \mathbf{X}_\rho^{-1} \mathbf{y}_\rho} \right) \\ &= \left(\frac{18}{14} \right) \left(\frac{15532.51 - 10068.46}{15906.4 - 15532.51} \right) \\ &= \left(\frac{18}{14} \right) \left(\frac{5464.05}{373.891} \right) \\ &= 1.07143 * 15505.58 \\ &= 16613.12 \end{aligned}$$

Cuadro 1. Análisis de varianza de *Diabrotica undecimpunctata* en términos de temperatura (x_1) y precipitación (x_2)

Fuente	DF	Suma de cuadrados	Cuadrado de la media	F-Valor	Pr>F
Modelo	6	5481.364194	913.560699	17.19	0.0034
Error	5	265.724973	53.144995		
Total	11	5747.089167			

Cuadro 2. Análisis de varianza de *Diabrotica balteata* en términos de temperatura (x_1) y precipitación (x_2)

Fuente	DF	Suma de cuadrados	Cuadrado de la media	F-Valor	Pr>F
Modelo	6	5258.695546	876.449258	40.58	0.0004
Error	5	107.993621	21.598724		
Total	11	5366.689167			

Cuadro 3. Análisis de varianza de *Diabrotica virgifera* en términos de temperatura (x_1) y precipitación (x_2)

Fuente	DF	Suma de cuadrados	Cuadrado de la media	F-Valor	Pr>F
Modelo	6	1.52425264	0.25404211	7.37	0.0223
Error	5	0.17241402	0.03448280		
Total	11	1.69666667			

Cuadro 4. Valores de la prueba de F, probabilidad P y coeficientes de determinación parciales R^2

Modelo	Valor de F	Probabilidad P	Coefficientes de determinación R^2
Y_1	40.58	0.0004	0.9798
Y_2	17.19	0.0034	0.9537
Y_3	7.37	0.0223	0.8983

Dado que el valor del estadístico fue de $W = 16613.12$ con probabilidad $P = 0^{**}$, podemos concluir que las curvas de crecimiento polinómicas poblacionales de las tres especies de *Diabrotica* son diferentes.

Al rechazar la prueba de igualdad de las tres curvas de crecimiento, se concluye que los sistemas de manejo más eficientes para estas plagas en el cultivo de maíz en Guanajuato, deben realizarse en diferente época (primavera, verano, otoño e invierno), considerando las condiciones de temperatura y precipitación bajo las cuales se presenta el máximo número de individuos adultos de cada una de las tres especies, ya que estas curvas de crecimiento, permiten estimar con buena precisión, el número de individuos adultos que se pueden presentar bajo ciertas condiciones de temperatura y precipitación.

El Cuadro 5 muestra los puntos críticos encontrados (x_1, x_2) , que establecen las condiciones climatológicas de temperatura y precipitación bajo las cuales se presenta la máxima densidad poblacional de adultos para cada una de las especies de *Diabrotica* estudiadas.

Los resultados obtenidos en esta investigación nos ayudan a comprender el crecimiento de las poblaciones de las tres especies de *Diabrotica*. Las diferencias entre especies nos ayudan al manejo integrado de estas plagas, la disminución en las pérdidas de los cultivos, la correcta absorción de nutrientes y la toma de decisiones. Por ello, es esencial que un modelo, lo más correctamente posible, describa lo que naturalmente está ocurriendo.

Cuadro 5. Puntos críticos de temperatura y precipitación (x_1, x_2) y valores máximos del número de adultos de tres especies de *Diabrotica* (Y)

Especie	Temperatura (T)	Precipitación (PP)	Número de adultos (Y)
<i>Diabrotica undecimpunctata</i>	14.0562	4.5000	3.2123
<i>Diabrotica balteata</i>	13.7000	168.0000	396.5033
<i>Diabrotica virgifera</i>	14.7020	4.5000	0.0837

Cabe resaltar que en lugar de comparar las medias (en diferentes tiempos o las medias de medidas repetidas), se procedió a comparar entre sí, los modelos polinómicos de las curvas de crecimiento poblacionales.

CONCLUSIONES Y RECOMENDACIONES

La prueba estadística para la hipótesis de igualdad de R modelos lineales o linealizables aplica cuando el interés sea la comparación de modelos en todo el recorrido de las variables independientes, en vez de la comparación de medias y/o varianzas, ya sean independientes o medidas repetidas.

Adicionalmente se tiene la ventaja de evaluar las tendencias, realizar predicciones para un punto dado y además se tiene la posibilidad de estimar los puntos críticos de las variables independientes en los cuales, las variables dependientes alcanzan su óptimo.

La solución propuesta para la prueba de igualdad de modelos puede ser generalizada al caso multivariado. Para tal objetivo se requerirá hacer un análisis similar al presentado en el trabajo, pero en tal caso, basado en la teoría del modelo lineal general multivariado y extender el teorema 1.3.

Otra posible generalización de interés es considerar que los modelos no son linealizables, en cuyo caso será necesario apoyarse en las técnicas de regresión no lineal univariadas o multivariadas, según el caso.

LITERATURA CITADA

- Díaz-García, J. A., y F. J. Caro-Lopera. 2006. Testing the equality of a set of multivariate linear models. *International Mathematical Forum*, 29:1423-1432.
- Draper, N. R. y H. Smith. 1981. *Applied regression analysis*. New York: Wiley.
- Fleetwood, K., F. Bursa y A. Yellowlees. 2015, Parallelism in practice: Approaches to parallelism in bioassays. *PDA J. Pharm. Sci. Technol.* 6:248-263.
- Graybill, F. A. 1976. *Matrices with applications in statistics (Second Edition Ed.)*. Belmont, California: Wadsworth International Group.
- Graybill, F. A. 2000. *Theory and application of the linear model*. Duxbury.
- Hodgson, J. A., C. D. Thomas, T. H. Oliver, B. J. Anderson, T. M. Brereton and E. E. Crone. 2011. Predicting insect phenology across space and time. *Global Change Biol.* 17:1289-1300.
- Kass, R. E., y A. E. Raftery. 1995. Bayes factor. *JASA*, 90(430):773-795.
- Martínez-Jaime, O. A., M. D. Salas-Araiza, J. A. Díaz-García, C. M. Bucio-Villalobos y E. Salazar-Solís. 2014. Comparison of population growth curves of adults of three species of *Diabrotica (Coleoptera: Chrysomelidae)* in maize (*Zea mays L.*) at Irapuato, Guanajuato, Mexico. *Southwest. Entomol.* 39 (3):581-593.
- Murtaugh, P. A., S. C. Emerson, P. B. Mcevoy y K. M. Higgs. 2012. The statistical analysis of insect phenology. *Environ. Entomol.* 41:355-361.
- Novick, S. J., H. Yang y J. J. Peterson. 2012. A bayesian approach to parallelism testing in bioassay. *Stat. Biopharm. Res.* 4:357-374.
- Pérez, D. J. F., M. L. L. García, Z. R. Álvarez y L. A. B. Rodríguez. 2010. "Diabroticas" como plagas del suelo. 361-372. En *Plagas del Suelo*. B. L. A. Rodríguez, and M. A. Morón, eds. Mundi-Prensa, México, D. F.
- Quesada-Moraga, E., P. Valverde-García y I. Garrido-Jurado. 2012. The effect of temperature and soil moisture on the development of the preimaginal Mediterranean fruit fly (*Diptera: Tephritidae*). *Environ. Entomol.* 41(4):966-970.
- Schwarz, G. 1978. Estimating the dimension of a model. *Ann. Statist.* 6 (2):461-464.
- Sin-nombre. (19 de 05 de 2018). Formas cuadráticas. Obtenido de Formas cuadráticas: <http://asignatura.us.es/amatiqui/php/activos/pdf/temas/Tema03.pdf>
- Soporte de Minitab 18. (02 de mayo de 2018). Soporte de Minitab 18. Obtenido de Soporte de Minitab 18: <https://support.minitab.com/es-mx/minitab/18/help->

and-how-to/probability-distributions-and-random-data/supporting-topics/distributions/chi-square-distribution/

APÉNDICE A

Elementos del Álgebra de Matrices

El concepto de matriz nos permite múltiples aplicaciones como lo son la manipulación, representación de datos y cálculo numérico y simbólico que se deriva de modelos matemáticos, los cuales son utilizados para resolver problemas en diferentes áreas del conocimiento.

Notación: en general las matrices serán denotadas con letras mayúsculas A, B, C, \dots , si A es una matriz, su componente (i, j) es A_{ij} o a_{ij} , así,

$$A = (A_{ij}) = (a_{ij}).$$

Si A consta de m filas y n columnas, A será denotada como $A \in \mathfrak{R}^{m \times n}$.

Si $A \in \mathfrak{R}^{m \times n}$, A se escribe por filas y columnas como

$$A = \begin{pmatrix} A'_{(1)} \\ A'_{(2)} \\ \vdots \\ A'_{(n)} \end{pmatrix} = (A_1, A_2, \dots, A_n),$$

$A_{(i)} \in \mathfrak{R}^n$
 $A_{(j)} \in \mathfrak{R}^m$.

La transpuesta de una matriz A , se denotará por A' y

$$(A')_{ij} = A_{ij}.$$

Si $A \in \mathfrak{R}^{m \times n}$, entonces $A' \in \mathfrak{R}^{n \times m}$.

I_n denotará la matriz identidad, $I_n \in \mathfrak{R}^{n \times n}$ cuando el contexto lo permita, se denotará por I . Similarmente 1_n denotará un vector de unos, $1_n \in \mathfrak{R}^{n \times 1} \equiv \mathfrak{R}^n$.

Para toda matriz $A \in \mathfrak{R}^{k \times k}$, $|A|$ o $\det(A)$ denotará el determinante de la matriz A .

Si A es una matriz cuadrada no singular ($|A| \neq 0$), A^{-1} denota la inversa de A .

Una matriz diagonal se denotará por D o $\text{diag}(d_i)$ o $\text{diag}(d_1, d_2, \dots, d_n)$ y es tal que $d_{ij} = 0$ si $i \neq j$.

Propiedades de la transpuesta y la inversa.

La transpuesta de una matriz es aquella obtenida mediante el intercambio de filas y columnas; la inversa de una matriz, es aquella cuyo valor multiplicado por la matriz original, nos da la matriz identidad.

TEOREMA A.1. Si $A \in \mathfrak{R}^{n \times n}$ tiene inversa, entonces A^{-1} es única. Más aún, $(A^{-1})^{-1} = A$.

TEOREMA A.2. Para matrices $A, B \in \mathfrak{R}^{n \times n}$, AB es no singular si y sólo si A y B son no singulares, más aún $(AB)^{-1} = B^{-1}A^{-1}$.

TEOREMA A.3. Si k es un escalar, $k \in \mathfrak{R}$ y A es no singular, entonces kA es no singular si $k \neq 0$ y $(kA)^{-1} = \frac{1}{k}A^{-1}$.

TEOREMA A.4. Sean A y B matrices del mismo orden y a, b escalares, entonces:

- i. $(aA)' = aA'$.
- ii. $(aA + bB)' = aA' + bB'$.

TEOREMA A.5. Suponga que el producto AB está definido, luego:

i. $(AB)' = B'A'$.

ii. $(A')' = A$.

iii. Si A es no singular, entonces $(A')^{-1} = (A^{-1})'$.

DEFINICIÓN A.1. (Matriz simétrica). Sea $A \in \mathfrak{R}^{n \times n}$ tal que $A = A'$. Se dice que A es una matriz simétrica.

TEOREMA A.6. Sea $A \in \mathfrak{R}^{m \times n}$ cualquier matriz, entonces $A'A$ y AA' son matrices simétricas.

TEOREMA A.7.

i. Para toda matriz cuadrada A , sea c_{ij} el cofactor del elemento a_{ij} , entonces:

$$|A| = \sum_{j=1}^n a_{ij}c_{ij} = \sum_{i=1}^n a_{ij}c_{ij}, \quad A \in \mathfrak{R}^{n \times n}.$$

ii. $|A| = |A'| = |A^{-1}|^{-1}$.

iii. Para matrices cuadradas $A, B \in \mathfrak{R}^{n \times n}$

$$|AB| = |A||B|.$$

DEFINICIÓN A.2. (Rango). El rango de una matriz A , $r(A)$, es el número de filas o columnas linealmente independientes. Equivalentemente, el rango de A es la dimensión de la submatriz de A de mayor orden, tal que $|A_1| \neq 0$, donde A_1 es tal submatriz. Más aún, se dice que $A \in \mathfrak{R}^{m \times n}$ es de rango completo si

$$r(\mathbf{A}) = \min [m, n].$$

TEOREMA A.8. *Sea \mathbf{C} cualquier matriz y \mathbf{A}, \mathbf{B} matrices de órdenes adecuados no singulares, entonces las siguientes matrices tienen el mismo rango*

$$\mathbf{C}, \mathbf{C}', \mathbf{AC}, \mathbf{CB}.$$

TEOREMA A.9. *Sean \mathbf{A} y \mathbf{B} matrices tales que el producto \mathbf{AB} está definido, entonces*

$$r(\mathbf{AB}) \leq [r(\mathbf{A}), r(\mathbf{B})].$$

Formas cuadráticas.

Una forma cuadrática es una aplicación que a cada vector le hace corresponder un número real de un polinomio; $\mathbf{A} \in \mathfrak{R}^{m \times m}$ es la matriz de la forma cuadrática y ésta es una matriz simétrica cuyos elementos de la diagonal principal son los coeficientes de los términos cuadráticos de la expresión polinómica, y los restantes elementos de la matriz son la mitad de los coeficientes de los términos no cuadráticos de dicha expresión (Sin-nombre, 2018)

DEFINICIÓN A.3. (Formas cuadráticas). *Sea $\mathbf{x} \in \mathfrak{R}^n$ una función $f: \mathfrak{R}^n \rightarrow \mathfrak{R}$ tal que*

$$f(\mathbf{x}) = \sum_{i,j=1}^n a_{ij}x_i x_j,$$

se llama la forma cuadrática.

Observe que si $\mathbf{A} = (a_{ij}), f(\mathbf{x})$ se puede reescribir como

$$f(x) = x'Ax,$$

donde A es llamada matriz de la forma cuadrática. Aún más, note que

$$f(x) = x'Ax = (x'Ax)' = x'A'x.$$

Pero note que si A no es simétrico, entonces

$$f(x) = \frac{x'Ax + x'A'x}{2} = x' \left(\frac{A + A'}{2} \right) x = x'Bx.$$

En donde $B = \left(\frac{A+A'}{2} \right)$ es tal que $B' = B$. Esto es, la matriz de una forma cuadrática f , siempre se puede seleccionar simétrica.

DEFINICIÓN A.4. (Matrices congruentes). *Dos matrices A y B son congruentes si y sólo si existe una matriz no singular C , tal que $B = C'AC$.*

TEOREMA A.10. *Si A y B son congruentes y A es simétrica, entonces B es simétrica.*

TEOREMA A.11. *Si A y B son matrices congruentes, $r(A) = r(B)$.*

DEFINICIÓN A.5. (Matriz ortogonal). *Sea $P \in \mathfrak{R}^{n \times n}$ se dice que P es una matriz ortogonal si y sólo si $P^{-1} = P'$.*

Algunas de las propiedades de las matrices ortogonales son resumidas a continuación.

TEOREMA A.12. *Sea P una matriz ortogonal $P = (P_1, P_2, \dots, P_n)$ entonces*

i. $P'P = PP' = I_n$.

- ii. $\|P_i\|^2 = 1$, donde $\|P_i\|^2 = P_i' P_i$.
- iii. $P_j' P_i = 0$ para todo $i \neq j$, $i, j = 1, 2, \dots, n$.
- iv. Para toda $A \in \mathfrak{R}^{n \times n}$, $|PAP'| = |P'AP| = |A|$.

Eigenvalores y eigenvectores.

(Alternativamente: eigenvalor= valor propio= valor característico= raíz característica= valor latente, etc., similarmente para eigenvector= vector propio, etc.)

Todo vector distinto de cero que satisfaga al sistema de ecuaciones lineales de la forma $Ax = \lambda x$, es un vector propio de A , asociado con el valor propio de λ ; de manera similar, los vectores propios se llaman vectores característicos (Kolman & Hill, 2006).

Considere la matriz $A \in \mathfrak{R}^{n \times n}$, se está interesado en resolver el sistema de ecuaciones lineales de la forma

$$Ax = \lambda x, \lambda \text{ es un escalar.}$$

DEFINICIÓN A.6. (Eigenvalor y eigenvector). Sea $A \in \mathfrak{R}^{n \times n}$ un eigenvalor de A es el único número λ tal que

$$Ax = \lambda x.$$

El vector $x \neq 0$ correspondiente, se conoce como eigenvector del eigenvalor λ .

TEOREMA A.13. Sea $A \in \mathfrak{R}^{n \times n}$. λ es un eigenvalor de A si y sólo si $P(\lambda) = |A - \lambda I_n|$ es tal que $P(\lambda) = 0$.

En el teorema 13 $P(\lambda) = |\mathbf{A} - \lambda \mathbf{I}_n|$ es llamado polinomio característico de \mathbf{A} y $P(\lambda) = 0$, ecuación característica.

Algunas propiedades de los eigenvalores son resumidas a continuación.

TEOREMA A. 14.

- i. Sea $\mathbf{A} \in \mathfrak{R}^{n \times n}$. \mathbf{A} tiene al menos un eigenvalor cero si y sólo si \mathbf{A} es singular.*
- ii. Sea $\mathbf{A} \in \mathfrak{R}^{n \times n}$ y \mathbf{C} cualquier matriz $\mathbf{C} \in \mathfrak{R}^{n \times n}$ no singular, entonces las matrices $\mathbf{A}, \mathbf{C}^{-1}\mathbf{A}\mathbf{C}, \mathbf{C}\mathbf{A}\mathbf{C}^{-1}$ todas tienen los mismos eigenvalores.*
- iii. \mathbf{A} y \mathbf{A}' tienen los mismos eigenvalores, aunque no necesariamente los mismos eigenvectores.*
- iv. Si λ es un eigenvalor de $\mathbf{A} \in \mathfrak{R}^{n \times n}$ y si \mathbf{x} es el correspondiente eigenvector del eigenvalor λ , entonces λ^k es un eigenvalor de \mathbf{A}^k y \mathbf{x} es un eigenvector de \mathbf{A}^k correspondiente a λ^k .*
- v. Si λ es un eigenvalor de \mathbf{A} , λ^{-1} es un eigenvalor de \mathbf{A}^{-1} .*
- vi. Si los eigenvalores de una matriz $\mathbf{A} \in \mathfrak{R}^{n \times n}$ son reales y distintos, entonces existe una matriz \mathbf{Q} no singular tal que $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{D}$, \mathbf{D} es una matriz diagonal.*
- vii. Sea $\mathbf{A} = \mathbf{A}' \in \mathfrak{R}^{n \times n}$, y además, sean λ_1 y λ_2 eigenvalores de \mathbf{A} y $\mathbf{x}_1, \mathbf{x}_2$ eigenvectores de \mathbf{A} , correspondientes a λ_1 y λ_2 , respectivamente. Si $\lambda_1 \neq \lambda_2$ entonces \mathbf{x}_1 y \mathbf{x}_2 son ortogonales.*
- viii. Sea $\mathbf{A} = \mathbf{A}' \in \mathfrak{R}^{n \times n}$ entonces existe \mathbf{P} ortogonal, tal que $\mathbf{P}'\mathbf{A}\mathbf{P} = \mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, λ_i son los eigenvalores de \mathbf{A} .*

Matrices particionadas.

La partición de una matriz es cuando se le divide en submatrices, utilizando líneas horizontales o verticales, éstas pueden particionarse en más de dos matrices y pueden ser operadas de acuerdo con las reglas generales de la operatoria de matrices.

Suponga que $B \in \mathfrak{R}^{n \times n}$, esto se puede escribir de la siguiente manera

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{matrix} m \\ n-m \\ m \\ n \end{matrix}$$

TEOREMA A.15.

i. Suponga que $|B_{11}| \neq 0$, entonces

$$|B| = |B_{11}| |B_{22} - B_{21} B_{11}^{-1} B_{12}| = |B_{11}| |B_{22 \cdot 1}|, \quad B_{22 \cdot 1} = B_{22} - B_{21} B_{11}^{-1} B_{12}.$$

ii. Similarmente si $|B_{22}| \neq 0$

$$|B| = |B_{22}| |B_{11} - B_{12} B_{22}^{-1} B_{21}| = |B_{22}| |B_{11 \cdot 2}|, \quad B_{11 \cdot 2} = B_{11} - B_{12} B_{22}^{-1} B_{21}.$$

iii. Si $A \in \mathfrak{R}^{n \times m}$ y $B \in \mathfrak{R}^{m \times n}$ entonces $|I_n - BA| = |I_m - AB|$.

iv. Si B es no singular y B_{11} y B_{22} también son no singulares, entonces si $A = B^{-1}$

$$A = \begin{pmatrix} B_{11}^{-1} & -B_{11}^{-1} B_{12} B_{22}^{-1} \\ -B_{22}^{-1} B_{21} B_{11}^{-1} & B_{22}^{-1} \end{pmatrix}.$$

Matrices no negativas.

Una matriz es no negativa cuando todos sus elementos son positivos o nulos.

DEFINICIÓN A.7. (Matrices no negativas). Sea $A = A' \in \mathfrak{R}^{n \times n}$, entonces:

- i. A es no negativa si y sólo si $x'Ax \geq 0$ para todo $x \in \mathfrak{R}^n$.
- ii. A es definida positiva, escribiendo $A > \mathbf{0}$, si y sólo si $x'Ax > 0$ para todo $x \in \mathfrak{R}^n$.
- iii. A es semidefinida positiva, denotando $A \geq \mathbf{0}$ si y sólo si $x'Ax \geq 0$ para todo $x \in \mathfrak{R}^n$ con la igualdad para algún $x \neq \mathbf{0}$.

Algunas de las propiedades de las matrices no negativas son:

TEOREMA A.16.

- i. Sea P una matriz no singular, entonces:
 - a. $A > \mathbf{0}$ si y sólo si $P'AP > \mathbf{0}$.
 - b. $A \geq \mathbf{0}$ si y sólo si $P'AP \geq \mathbf{0}$.
- ii. $A > \mathbf{0}$ si y sólo si todos sus eigenvalores, digamos $\lambda_i, i = 1, 2, \dots, n$, son positivos. $A \geq \mathbf{0}$ si y sólo si todos sus eigenvalores λ_i son no negativos y al menos un eigenvalor es cero.
- iii. Sea $C \in \mathfrak{R}^{n \times n}$, entonces:
 - a. CC' es una matriz no negativa de orden $n \times n$.
 - b. $C'C$ es una matriz no negativa de orden $k \times k$.
 - c. Si $r(C) = n$, entonces $CC' > \mathbf{0}$ y si $r(C) < n$, entonces $CC' \geq \mathbf{0}$.
 - d. Si $r(C) = k$, entonces $C'C > \mathbf{0}$ y si $r(C) < k$, entonces $C'C \geq \mathbf{0}$.
- iv. Si $A = A' \in \mathfrak{R}^{n \times n}$, existe $C \in \mathfrak{R}^{n \times n}$ tal que $A = C'C = (CC')$.

Inversa generalizada o inversa de Moore-Penrose.

Esta se refiere al ajuste óptimo de un sistema de ecuaciones lineales que no posee solución única, y las inversas generalizadas nos permiten caracterizar soluciones aproximadas.

DEFINICIÓN A.8. (Inversa de Moore-Penrose). Si $A \in \mathfrak{R}^{m \times n}$, entonces existe A^- única matriz, $A^- \in \mathfrak{R}^{n \times m}$ tal que

- i. $(AA^-) = (AA^-)'$,
- ii. $A^-A = (A^-A)'$,
- iii. $AA^-A = A$,
- iv. $A^-AA^- = A^-$.

A continuación se resumen algunas de sus propiedades.

TEOREMA A.17.

- i. Para toda matriz A , $(A^-)^- = A$ y $A'^- = (A^-)'$.
- ii. $r(A) = r(AA^-) = r(A^-A) = r(A^-)$.
- iii. Si $A = A'$, entonces $A^- = (A^-)'$.
- iv. Si $A = A' = A^2$, entonces $A^- = A$.
- v. Las siguientes matrices son simétricas ($A = A'$) e idempotentes ($A = A^2$), A^-A , $I - A^-A$, AA^- y $I - AA^-$.
- vi. $(A'A)^- = A^-A'^- = A^-A'^-$.
- vii. Sean $A \in \mathfrak{R}^{n \times m}$ y $B \in \mathfrak{R}^{m \times n}$, entonces $AB = AA^-$ si y sólo si
 - a. $AB = (AB)'$,

b. $ABA = A$.

Inverso condicional.

DEFINICIÓN A.9. (Inverso condicional). Sea $A \in \mathfrak{R}^{m \times n}$ una matriz y sea $A^c \in \mathfrak{R}^{n \times m}$, se dice que es un inverso condicional de A si y sólo si $AA^cA = A$.

Algunas propiedades básicas del inverso condicional son presentadas en el siguiente resultado.

TEOREMA A.18.

- i. El inverso de Moore-Penrose es un inverso condicional, pero no viceversa.
- ii. Sea $X \in \mathfrak{R}^{m \times n}$ tal que $r(X) = r > 0$, entonces:
 - a. $r(X^c) \geq 0$.
 - b. X^cX y XX^c son idempotentes.
 - c. $r(X^cX) = r(XX^c) = r(X) = r$.
 - d. $X^cX = I_n$ si $r(X) = n$.
 - e. $XX^c = I_m$ si $r(X) = m$.
 - f. Sea $X \in \mathfrak{R}^{n \times n}$, su traza denotada como $\text{tr}(A)$ se define como

$$\text{tr}(A) = \sum_{i=1}^n a_{ii},$$

luego $\text{tr}(XX^c) = \text{tr}(X^cX) = r$.

- g. Si X^c es un inverso condicional de X , entonces $(X^c)'$ es un inverso condicional de X' .

- h. Sea $K = X(X'X)^c X'$ entonces K es invariante bajo todo inverso condicional de $X'X$.
- i. $K = X(X'X)^c X' = XX^-$.
- j. $K = K' = K^2$.
- k. $r(K) = \text{tr}(K) = r(X)$.
- l. $KX = X$ y $X'K = X'$.
- m. $(X'X)^c X'$ es un inverso condicional de X y $X(X'X)^c$ es un inverso condicional de X' .

Sistema de ecuaciones lineales.

Las matrices nos permiten saber de manera rápida, si un sistema de ecuaciones tiene soluciones y que tipo; un sistema puede no tener solución, tener una única solución o tener infinitas soluciones.

Sea $A \in \mathfrak{R}^{n \times m}$ y considere el sistema de ecuaciones lineales

$$Ax = g, \quad x \in \mathfrak{R}^m, \quad g \in \mathfrak{R}^n$$

El sistema tiene solución, al menos una sí y sólo si g pertenece al espacio de columnas de A , $g \in \mathfrak{C}_c(A)$. Equivalentemente:

TEOREMA A.19. Si $A \in \mathfrak{R}^{n \times n}$ es no singular, entonces $Ax = g$ tiene solución única para todo $g \in \mathfrak{R}^n$, más aún, una solución única está dada por $x = A^{-1}g$.

TEOREMA A.20. El sistema de ecuaciones lineales $Ax = g$ tiene solución (al menos una) si y sólo si

i. $r(A) = r(A : g)$ o $AA^c g = g$ o $AA^- g = g$.

- ii. Tiene solución única para todo $\mathbf{g} \in \mathfrak{R}^m$ si $r(\mathbf{A}) = m$, $\mathbf{A} \in \mathfrak{R}^{m \times n}$.
- iii. Si \mathbf{A}^c es un inverso condicional de \mathbf{A} , toda solución del sistema se obtiene variando \mathbf{h} , donde $\mathbf{h} \in \mathfrak{R}^{m \times n}$, en

$$\mathbf{x}_0 = \mathbf{A}^c \mathbf{g} + (\mathbf{I} - \mathbf{A}^c \mathbf{A}) \mathbf{h}.$$

- iv. Si un sistema de ecuaciones tiene solución se dice que es consistente y sus soluciones se obtienen alternativamente como

$$\mathbf{x}_0 = \mathbf{A}^- \mathbf{g} + (\mathbf{I} - \mathbf{A}^- \mathbf{A}) \mathbf{h},$$

la cual es única si y sólo si $\mathbf{A}^c \mathbf{A} = \mathbf{A}^- \mathbf{A} = \mathbf{I}$.

Nota: Recuerde que dado un vector, su norma se define y denota como:

Sea $\mathbf{x} \in \mathfrak{R}^n$

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}' \mathbf{x}}$$

es tal que

- i. $\|\mathbf{x}\| = 0$ si y sólo si $\mathbf{x} = \mathbf{0}$.
- ii. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.
- iii. Para $\alpha \in \mathfrak{R}$ $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$, donde $|a| = \text{abs}(a)$.

Por el teorema 16, note que

$$\mathbf{x}' \mathbf{A} \mathbf{x} = \mathbf{x}' \mathbf{C}' \mathbf{C} \mathbf{x} = (\mathbf{C} \mathbf{x})' \mathbf{C} \mathbf{x} = \|\mathbf{C} \mathbf{x}\|^2, \quad \mathbf{A} = \mathbf{A}' = \mathbf{C}' \mathbf{C}.$$

APÉNDICE B

Distribución Normal Multivariada

Es una generalización de la distribución normal unidimensional a dimensiones superiores. Los datos multivariados son el resultado de observar un vector aleatorio cuyos componentes son variables aleatorias definidas sobre el mismo espacio de probabilidad.

DEFINICIÓN B.1. (Distribución normal univariada). *Se dice que una variable aleatoria x tiene una distribución normal si su función de densidad está dada por*

$$\eta(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2}\right), \quad x \in \mathfrak{R}, \text{ (ver Apéndice A)}$$

donde los parámetros μ y σ^2 son tales que $\mu \in \mathfrak{R}$ y $\sigma^2 > 0$, escribiendo $x \sim N(\mu, \sigma^2)$.

Algunas de sus propiedades son:

TEOREMA B.1. *Suponga que $x \sim N(\mu, \sigma^2)$, entonces*

- i. $\varphi_x(t) = E(\exp(itx)) = \exp\left(it\mu - \frac{1}{2}t^2\sigma^2\right)$, $i = \sqrt{-1}$ (función característica).
- ii. $E(x) = \mu$.
- iii. $\text{var}(x) = \sigma^2$.
- iv. Si $x_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2, \dots, n$ entonces para a_i constantes, $i = 1, 2, \dots, n$,

$$\sum_{i=1}^n a_i x_i \sim N\left(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2\right), \quad x_i \text{ independientes.}$$

DEFINICIÓN B.2. (Distribución normal estándar multivariada). Sea $\mathbf{z} \in \mathfrak{R}^k$ un vector aleatorio (ver Apéndice A), se dice que \mathbf{z} tiene una distribución normal estándar k -dimensional si

- i. Z_1, Z_2, \dots, Z_n son variables aleatorias independientes.
- ii. $Z_i \sim N(0,1)$ para toda $i = 1, 2, \dots, k$.

por lo tanto

- a. $\varphi_{\mathbf{z}}(\mathbf{t}) = E(\exp(i\mathbf{t}'\mathbf{x})) = \exp\left(-\frac{1}{2}\|\mathbf{t}\|^2\right)$.
- b. $\eta(\mathbf{z}; \mathbf{0}, \mathbf{I}_k) = \frac{1}{(2\pi)^{k/2}} \exp\left(-\frac{1}{2}\|\mathbf{z}\|^2\right)$.
- c. $E(\mathbf{z}) = \mathbf{0}$.
- d. $\text{Cov}(\mathbf{z}) = \mathbf{I}_k$.
- e. Sea $\mathbf{\Gamma}$ cualquier matriz, $\mathbf{\Gamma} \in \mathfrak{R}^{p \times k}$, $\boldsymbol{\gamma} \in \mathfrak{R}^p$ y $\mathbf{z} \sim N_k(\mathbf{0}, \mathbf{I}_k)$ y defina $\mathbf{x} = \mathbf{\Gamma}\mathbf{z} + \boldsymbol{\gamma}$, entonces
 - 1°. $E(\mathbf{x}) = \boldsymbol{\gamma}$.
 - 2°. $\text{Cov}(\mathbf{x}) = \mathbf{\Gamma}\mathbf{\Gamma}'$.
 - 3°. $\varphi_{\mathbf{x}}(\mathbf{t}) = E(\exp(i\mathbf{t}'\mathbf{x})) = \exp\left(i\mathbf{t}'\boldsymbol{\gamma} - \frac{1}{2}\|\mathbf{\Gamma}'\mathbf{t}\|^2\right)$.

DEFINICIÓN B.3. (Distribución normal multivariada). Sea $\mathbf{x} \in \mathfrak{R}^p$ un vector aleatorio, \mathbf{x} tiene una distribución normal p -dimensional si y sólo si $\boldsymbol{\mu} \in \mathfrak{R}^p$ y $\boldsymbol{\Sigma} \in \mathfrak{R}^{p \times p}$, $\boldsymbol{\Sigma} \geq 0$, son tales que

$$\varphi_{\mathbf{x}}(\mathbf{t}) = \exp\left(i\mathbf{t}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}\right),$$

escribiendo $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, más aún, si $\boldsymbol{\Sigma} > 0$, entonces

$$\eta(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{n/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right),$$

y $E(\mathbf{x}) = \boldsymbol{\mu}$ y $\text{Cov}(\mathbf{x}) = \boldsymbol{\Sigma}$.

TEOREMA B.2. *Suponga que $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ entonces*

i. *Para cualquier $\mathbf{b} \in \mathfrak{R}^n$ y $\mathbf{A} \in \mathfrak{R}^{n \times p}$*

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b} \sim N_n(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}').$$

ii. *Considere las siguientes particiones*

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \begin{matrix} p_1 x_1 \\ p_2 x_1 \end{matrix} \quad \mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \begin{matrix} p_1 x_1 \\ p_2 x_1 \end{matrix} \quad \text{y} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \begin{matrix} p_1 \\ p_2 \end{matrix} \quad p_1 + p_2 = p.$$

Entonces $\mathbf{x}_i \sim N_{p_i}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_{ii})$, $i = 1, 2, \dots, n$

iii. *\mathbf{x}_1 y \mathbf{x}_2 son independientes si y sólo si $\boldsymbol{\Sigma}_{12} = \text{Cov}(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{0}$.*

iv. *$\mathbf{x}_1 / \mathbf{x}_2 = \mathbf{C}_2 \sim N_{p_i}(\boldsymbol{\mu}_i - \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} (\boldsymbol{\mu}_2 - \mathbf{C}_2), \boldsymbol{\Sigma}_{11 \cdot 2})$, donde*

$$\boldsymbol{\Sigma}_{11 \cdot 2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\Sigma}_{21}, \quad \boldsymbol{\Sigma}_{22} > \mathbf{0}.$$

Distribución de Formas Cuadráticas

En esta sección se exponen brevemente algunas distribuciones relacionadas con la distribución normal multivariada y las formas cuadráticas.

Distribución Ji-Cuadrada.

La distribución Ji-cuadrada es una distribución continua que se especifica por los grados de libertad y el parámetro de no centralidad. La distribución es positivamente asimétrica, pero la asimetría disminuye al aumentar los grados de libertad (Soporte de Minitab 18, 2018)

DEFINICIÓN B.4. (Distribución Ji-Cuadrada). *Y tiene una distribución χ^2 (Ji-Cuadrada) con n grados de libertad, escribiendo $Y \sim \chi_n^2$ si su densidad es*

$$f_Y(y) = \frac{1}{2^{n/2} \Gamma[n/2]} \exp\left(-\frac{1}{2} y\right) y^{n/2-1}, \quad y > 0.$$

Tal densidad se denotará como $\chi^2(Y; n)$.

Algunas de sus propiedades son

TEOREMA B.3. *Asuma que $Y \sim \chi_n^2$, entonces*

- i. $\varphi_Y(t) = (1 - 2it)^{-n/2}$.
- ii. $E(Y) = n$.
- iii. $\text{Var}(Y) = 2n$.
- iv. Si $W_i \sim \chi_{n_i}^2, i = 1, 2, \dots, r$ independientes,

$$\sum_{i=1}^r W_i \sim \chi_{\sum_{i=1}^r n_i}^2.$$

v. Suponga que $\mathbf{z} \sim N_k(\mathbf{0}, \mathbf{I})$ entonces

$$\mathbf{z}'\mathbf{z} = \|\mathbf{z}\|^2 = \sum_{i=1}^k Z_i^2 \sim \chi_k^2.$$

Distribución Ji-Cuadrada no centrada.

Suponga que $\mathbf{y} \sim N_n(\boldsymbol{\mu}, \mathbf{I})$, se está interesado en la distribución de la forma cuadrática

$$\mathbf{y}'\mathbf{y} = \|\mathbf{y}\|^2 = \sum_{i=1}^n Y_i^2,$$

tal distribución es llamada distribución Ji-Cuadrada no centrada.

TEOREMA B.4. Suponga que $\mathbf{y} \sim N_n(\boldsymbol{\mu}, \mathbf{I}_n)$ y defina $V = \|\mathbf{y}\|^2 = \mathbf{y}'\mathbf{y}$ entonces

- i. $\varphi_V(t) = (1 - 2it)^{-n/2} \exp\left(\frac{2it\lambda}{(1-2it)}\right), \lambda = \frac{1}{2}\|\boldsymbol{\mu}\|^2 = \frac{1}{2}\boldsymbol{\mu}'\boldsymbol{\mu}.$
- ii. Su función de densidad está dada por

$$f_V(v) = \sum_{k=0}^{\infty} \frac{\exp(-\lambda) \lambda^k}{k!} \chi^2(v; n + 2k), \quad v > 0$$

Donde $\chi_v^2(n + 2k)$ es la densidad de $v \sim \chi_{n+2k}^2$.

- iii. $E(V) = n + 2\lambda.$
- iv. $\text{Var}(V) = 2(n + 4\lambda).$

Así, la distribución de $V = \mathbf{y}'\mathbf{y} = \|\mathbf{y}\|^2$ es llamada distribución Ji-Cuadrada no centrada con n grados de libertad y parámetro de no centralidad λ , denotando este hecho como

$$V \sim \chi_{n,\lambda}^2.$$

Bajo condiciones más generales se tienen los siguientes resultados.

TEOREMA B.5. *Suponga que $\mathbf{y} \sim N_n(\boldsymbol{\mu}, \mathbf{I})$ y sea $u = \mathbf{y}' \mathbf{A} \mathbf{y}$, entonces*

$$u \sim \chi_{r(\mathbf{A}),\lambda}^2, \lambda = \frac{1}{2} \boldsymbol{\mu}' \mathbf{A} \boldsymbol{\mu},$$

si y sólo si $\mathbf{A} = \mathbf{A}^2$ (es idempotente) (ver Apéndice A).

Además,

TEOREMA B.6. *Sea $\boldsymbol{\Sigma} > 0$ y suponga que $\mathbf{y} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ además, defina $u = \mathbf{y}' \mathbf{A} \mathbf{y}$, entonces u tiene una distribución Ji-Cuadrada no centrada, $u \sim \chi_{r(\mathbf{A}),\lambda}^2$, $\lambda = \frac{1}{2} \boldsymbol{\mu}' \mathbf{A} \boldsymbol{\mu}$ si y sólo si $\boldsymbol{\Sigma}$ es un inverso condicional de \mathbf{A} , es decir, si $\mathbf{A} \boldsymbol{\Sigma} \mathbf{A} = \mathbf{A}$.*

Finalmente

TEOREMA B.7. *Suponga que \mathbf{y} es un vector aleatorio tal que $E(\mathbf{y}) = \boldsymbol{\mu}$ y $\text{Cov}(\mathbf{y}) = \boldsymbol{\Sigma}$, entonces*

$$E(\mathbf{y}' \mathbf{A} \mathbf{y}) = \text{tr} \mathbf{A} \boldsymbol{\Sigma} + \boldsymbol{\mu}' \mathbf{A} \boldsymbol{\mu}, \quad \mathbf{A} \geq \mathbf{0}.$$

Independencia de formas lineales y cuadráticas.

Asuma que $\mathbf{y} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y se tienen las formas cuadráticas $\mathbf{y}' \mathbf{A} \mathbf{y}$ y $\mathbf{y}' \mathbf{B} \mathbf{y}$ (o la forma lineal $\mathbf{B} \mathbf{y}$) se está interesado en saber cuándo éstas son independientes.

TEOREMA B.8. *Suponga que $\mathbf{y} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y sean $\mathbf{A} \in \mathfrak{R}^{n \times n}$, $\mathbf{A} \geq \mathbf{0}$ y $\mathbf{B} \in \mathfrak{R}^{p \times n}$, si $\mathbf{B} \boldsymbol{\Sigma} \mathbf{A} = \mathbf{0}$, entonces $\mathbf{B} \mathbf{y}$ y $\mathbf{y}' \mathbf{A} \mathbf{y}$ son independientes.*

Similarmente

TEOREMA B.9. *Asuma que $y \sim N_n(\mu, \Sigma)$ y sean $A \geq 0$ y $B \geq 0$, $A, B \in \mathfrak{R}^{n \times n}$, entonces sí $B\Sigma A = 0$, $y' Ay$ y $y' By$ son independientes (ver Apéndice A).*

Distribución F no centrada.

Recuerde que si u_1 y u_2 son variables aleatorias independientes tales que $u_1 \sim \chi_{n_1}^2$ y $u_2 \sim \chi_{n_2}^2$ la variable definida como

$$W = \frac{\frac{u_1}{n_1}}{\frac{u_2}{n_2}} = \frac{n_2 u_1}{n_1 u_2},$$

se dice que se tiene una distribución F y se denotará escribiendo $W \sim F_{n_1, n_2}$ y su correspondiente densidad por $F(W; n_1, n_2)$, más aún

$$F(W; n_1, n_2) = \frac{\Gamma[(n_1 + n_2)/2]}{\Gamma[n_1/2]\Gamma[n_2/2]} \left(\frac{n_1}{n_2}\right)^{n_1/2} \frac{W^{(n_1-2)/2}}{(1 + (n_1/n_2)W)^{(n_1+n_2)/2}},$$

donde $W > 0$, $n_1, n_2 \in \mathbb{N}$,

- i. $E(W) = \frac{n_2}{n_2-2}, n_2 > 2,$
- ii. $\text{Var}(W) = \frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}, n_2 > 4.$

Análogamente se está interesado en la distribución de W cuando $u_1 \sim \chi_{n_1, \lambda}^2$, en cuyo caso se dice que W tiene una distribución F no centrada con n_1 y n_2 grados de libertad y parámetro de no centralidad λ , escribiendo

$$W \sim F_{n_1, n_2, \lambda},$$

y su correspondiente densidad será denotada como $F(W; n_1, n_2, \lambda)$.

TEOREMA B.10. Si $W \sim F_{n_1, n_2, \lambda}$ entonces

$$F(W; n_1, n_2, \lambda) = \sum_{k=0}^{\infty} \frac{\exp(-\lambda) \lambda^k}{k!} F(s; n_1 + 2k, n_2), \quad s = \frac{n_1}{n_1 + 2k} W, \quad W > 0.$$

Nota: Observe que cuando $\lambda = 0$, la densidad F no centrada se reduce a la densidad F centrada, esto es

$$F(W; n_1, n_2, 0) \equiv F(W; n_1, n_2).$$

Finalmente observe que cuando $W \sim F_{n_1, n_2, \lambda}$

$$E(W) = \frac{n_2(n_1 + \lambda)}{n_1(n_2 - 2)}, \quad n_2 > 2,$$

y

$$\text{Var}(W) = 2 \left(\frac{n_1}{n_2} \right)^2 \left[\frac{(n_1 + \lambda)^2 + (n_1 + 4\lambda)(n_2 - 2)}{(n_2 - 2)^2(n_2 - 4)} \right], \quad n_2 > 4.$$

APÉNDICE C

Modelo Lineal General

DEFINICIÓN C.1. (Modelo lineal general). Sea $\mathbf{y} \in \mathfrak{R}^n$ un vector aleatorio; sean además $\mathbf{X} \in \mathfrak{R}^{n \times p}$, una matriz constante; $\boldsymbol{\beta} \in \mathfrak{R}^p$ vector fijo pero desconocido y $\boldsymbol{\epsilon} \in \mathfrak{R}^n$ un vector aleatorio no observable, finalmente suponga que

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

La relación anterior define un modelo lineal general, donde se supone que $E(\boldsymbol{\epsilon}) = \mathbf{0}$ y la matriz de covarianzas es $\boldsymbol{\Sigma}$, $\text{Cov}(\mathbf{y}) = \boldsymbol{\Sigma}$.

Nota: este modelo contiene muchos casos especiales dependiendo de:

- (1) La distribución de $\boldsymbol{\epsilon}$,
- (2) La estructura de la matriz de varianzas y covarianzas $\boldsymbol{\Sigma}$, y
- (3) El rango y la estructura de \mathbf{X} .

Para específicas condiciones (1)-(3), este modelo contiene como casos particulares los siguientes modelos:

- a) Modelo lineal general (x fija observable),
- b) Modelo de regresión lineal (y ; x aleatorios con $f(x_1, x_2, \dots, x_n)$),
- c) Modelos lineales y de regresión linealizables,
- d) Modelos de regresión de efectos mixtos ($\boldsymbol{\beta}$; aleatorios y/o fijos independientes de $\boldsymbol{\epsilon}$),
- e) Modelos de regresión estocástica (x ; aleatoria independiente de $\boldsymbol{\epsilon}$),

- f) Modelos de regresión con errores en las variables (x_{ij} 's, y_{ij} 's; no observables aleatorios o no),
- g) Modelos de diseño (modelo cualitativo),
- h) Modelos de componentes de varianza (modelo cualitativo).

Incluso los modelos de series de tiempo pueden ser establecidos con la forma de un modelo lineal general.

Si en el modelo anterior $r(\mathbf{X}) = p$, éste es llamado modelo lineal general de rango completo, hecho que será asumido en esta sección.

En relación a la distribución del vector aleatorio $\boldsymbol{\varepsilon}$, hay dos posibilidades.

Caso I. $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2 \mathbf{I})$ o equivalentemente $\mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$ (ver Apéndice B).

Caso II. $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ y $\text{Cov}(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}_n$ equivalentemente, el espacio paramétrico está definido como

$$\Omega = \left\{ \begin{pmatrix} \boldsymbol{\beta} \\ \sigma^2 \end{pmatrix} \mid \boldsymbol{\beta} \in \mathbb{R}^p, \sigma^2 > 0 \right\}.$$

Caso I.

Estimación de $\boldsymbol{\beta}$ y σ^2 .

Aplicando el método de verosimilitud máxima, se tiene que $\mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I}_n)$, luego la función de verosimilitud está dada como:

$$\mathcal{L}(\boldsymbol{\beta}, \sigma^2; \mathbf{y}) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2\right),$$

aplicando logaritmos se tiene:

$$\Omega(\boldsymbol{\beta}, \sigma^2; \mathbf{y}) = \log \mathcal{L}(\boldsymbol{\beta}, \sigma^2; \mathbf{y}),$$

$$\Omega(\boldsymbol{\beta}, \sigma^2; \mathbf{y}) = -\frac{n}{2} \log \sigma^2 - \frac{n}{2} \log 2\pi - \frac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2,$$

sacando el diferencial se obtiene

$$d\Omega = -\frac{n}{2\sigma^2} d\sigma^2 - \frac{(-1)(\sigma^2)^2}{2} d\sigma^2 \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 - \frac{1}{2\sigma^2} d\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2.$$

Note que

$$\begin{aligned} d\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 &= d(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = d(\mathbf{y}'\mathbf{y} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}) \\ &= d(\mathbf{y}'\mathbf{y} - 2\mathbf{y}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}) \\ &= -2\mathbf{y}'\mathbf{X}d\boldsymbol{\beta} + d\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}d\boldsymbol{\beta} \\ &= -2\mathbf{y}'\mathbf{X}'d\boldsymbol{\beta} + 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}d\boldsymbol{\beta} \\ &= (-2\mathbf{y}'\mathbf{X} + 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{X})d\boldsymbol{\beta}, \end{aligned}$$

de donde

$$d\Omega = \left(-\frac{n}{2\sigma^2} + \frac{\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2}{2(\sigma^2)^2} \right) d\sigma^2 + \frac{1}{\sigma^2} (\mathbf{y}'\mathbf{X}' - \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}) d\boldsymbol{\beta}.$$

Luego

$$\frac{\partial \Omega}{\partial \boldsymbol{\beta}} = \left[\frac{1}{\sigma^2} (\mathbf{y}'\mathbf{X}' - \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}) \right]' = \frac{1}{\sigma^2} (\mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\boldsymbol{\beta}),$$

y

$$\frac{\partial \Omega}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2}{2(\sigma^2)^2}.$$

Igualando a cero las derivadas y denotando la solución del sistema restante por $\tilde{\boldsymbol{\beta}}$ y $\tilde{\sigma}^2$ se obtiene

$$0 = \frac{\partial \mathcal{L}}{\partial \boldsymbol{\beta}} = \frac{1}{\tilde{\sigma}^2} [\mathbf{X}' \mathbf{y} - \mathbf{X}' \mathbf{X} \tilde{\boldsymbol{\beta}}]$$

$$\Rightarrow \mathbf{X}' \mathbf{y} - \mathbf{X}' \mathbf{X} \tilde{\boldsymbol{\beta}} = \mathbf{0}$$

$$\Rightarrow \mathbf{X}' \mathbf{X} \tilde{\boldsymbol{\beta}} = \mathbf{X}' \mathbf{y}.$$

Además

$$0 = \frac{\partial \mathcal{L}}{\partial \sigma^2} = -\frac{n}{2\tilde{\sigma}^2} + \frac{\|\mathbf{y} - \mathbf{X} \tilde{\boldsymbol{\beta}}\|^2}{2(\tilde{\sigma}^2)^2}$$

$$\Rightarrow \frac{n}{2\tilde{\sigma}^2} = \frac{\|\mathbf{y} - \mathbf{X} \tilde{\boldsymbol{\beta}}\|^2}{2(\tilde{\sigma}^2)^2}$$

$$\Rightarrow \frac{2(\tilde{\sigma}^2)^2}{2\tilde{\sigma}^2} = \frac{1}{n} \|\mathbf{y} - \mathbf{X} \tilde{\boldsymbol{\beta}}\|^2$$

$$\Rightarrow \tilde{\sigma}^2 = \frac{1}{n} \|\mathbf{y} - \mathbf{X} \tilde{\boldsymbol{\beta}}\|^2.$$

Diferenciando nuevamente con respecto a $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\beta} \\ \sigma^2 \end{pmatrix}$ se tiene:

$$\begin{aligned} d^2 \mathcal{L} = & -\frac{n}{2} (-1)(\sigma^2)^{-2} d^2 \sigma^2 + \frac{(-2)(\sigma^2)^{-3} d^2 \sigma^2}{2} \|\mathbf{y} - \mathbf{X} \boldsymbol{\beta}\|^2 \\ & + (-1)(\sigma^2)^{-2} d\sigma^2 [(\mathbf{y} - \mathbf{X} \boldsymbol{\beta})' \mathbf{X}] d\boldsymbol{\beta} \\ & + \frac{d\sigma^2}{2(\sigma^2)^2} [-2(\mathbf{y} - \mathbf{X} \boldsymbol{\beta})' \mathbf{X} d\boldsymbol{\beta}] + \frac{1}{\sigma^2} (-d\boldsymbol{\beta}' \mathbf{X}' \mathbf{X} d\boldsymbol{\beta}), \end{aligned}$$

$$\begin{aligned}
&= \frac{nd^2\sigma^2}{2(\sigma^2)^2} - \frac{\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2}{(\sigma^2)^3} d^2\sigma^2 - \frac{d\sigma^2}{(\sigma^2)^2} ((\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{X}) d\boldsymbol{\beta} \\
&\quad - \frac{d\sigma^2}{(\sigma^2)^2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{X} d\boldsymbol{\beta} - \frac{1}{\sigma^2} d\boldsymbol{\beta}' \mathbf{X}' \mathbf{X} d\boldsymbol{\beta}, \\
&= \left[\frac{n}{2(\sigma^2)^2} - \frac{\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2}{(\sigma^2)^3} \right] d^2\sigma^2 + d\sigma^2 \left[-\frac{2(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{X}}{(\sigma^2)^2} \right] + d\boldsymbol{\beta}' \left[-\frac{\mathbf{X}' \mathbf{X}}{\sigma^2} \right] d\boldsymbol{\beta}.
\end{aligned}$$

De donde dada la simetría de $\frac{\partial^2 \mathcal{L}}{\partial(\sigma^2)^2}$ y $\frac{\partial^2 \mathcal{L}}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}$, se tiene que

$$\begin{aligned}
\frac{\partial^2 \mathcal{L}}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} &= \begin{pmatrix} \frac{\partial^2 \mathcal{L}}{\partial(\sigma^2)^2} & \frac{\partial^2 \mathcal{L}}{\partial\sigma^2\partial\boldsymbol{\beta}'} \\ \frac{\partial^2 \mathcal{L}}{\partial\boldsymbol{\beta}\partial\sigma^2} & \frac{\partial^2 \mathcal{L}}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'} \end{pmatrix}, \\
&= \begin{pmatrix} \frac{n}{2(\sigma^2)^2} - \frac{\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2}{(\sigma^2)^3} & -\frac{2(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{X}}{(\sigma^2)^2} \\ -\frac{2(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{X}}{(\sigma^2)^2} & -\frac{\mathbf{X}' \mathbf{X}}{\sigma^2} \end{pmatrix}.
\end{aligned}$$

Evaluando en $\tilde{\boldsymbol{\theta}} = \begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\sigma}^2 \end{pmatrix}$ se tiene:

i.

$$\begin{aligned}
\frac{n}{2(\tilde{\sigma}^2)^2} - \frac{\|\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}\|^2}{(\tilde{\sigma}^2)^3} &= \frac{n}{2(\tilde{\sigma}^2)^2} - \frac{n}{(\tilde{\sigma}^2)^3} \left(\frac{1}{n} \|\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}\|^2 \right) = \frac{n}{2(\tilde{\sigma}^2)^2} - \frac{n}{(\tilde{\sigma}^2)^3} \tilde{\sigma}^2 \\
&= \frac{n}{2(\tilde{\sigma}^2)^2} - \frac{n}{(\tilde{\sigma}^2)^2} = -\frac{n}{2(\tilde{\sigma}^2)^2}.
\end{aligned}$$

ii.

$$-\frac{2}{(\tilde{\sigma}^2)^2} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})' \mathbf{X} = -\frac{2}{(\tilde{\sigma}^2)^2} (\mathbf{y}' \mathbf{X} - \tilde{\boldsymbol{\beta}}' \mathbf{X}' \mathbf{X}) = 0.$$

De donde

$$\frac{\partial^2 \mathcal{L}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \Big|_{\tilde{\boldsymbol{\theta}}} = \begin{pmatrix} -\frac{n}{2(\tilde{\sigma}^2)^2} & 0 \\ 0 & -\frac{\mathbf{X}'\mathbf{X}}{\tilde{\sigma}^2} \end{pmatrix} = - \begin{pmatrix} \frac{n}{2(\tilde{\sigma}^2)^2} & 0 \\ 0 & \frac{\mathbf{X}'\mathbf{X}}{\tilde{\sigma}^2} \end{pmatrix} < 0.$$

Como $\mathbf{X} \in \mathfrak{R}^{n \times p}$, $r(\mathbf{X}) = p$, se tiene que $r(\mathbf{X}'\mathbf{X}) = r(\mathbf{X}) = p$, así $\mathbf{X}'\mathbf{X} \in \mathfrak{R}^{p \times p}$ es no singular, de donde $\mathbf{X}'\mathbf{X} > \mathbf{0}$, más aún el sistema $\mathbf{X}'\mathbf{X}\tilde{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$ tiene solución única dada por

$$\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} = \mathbf{X}^- \mathbf{y}.$$

Dado que $\frac{\partial^2 \mathcal{L}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \Big|_{\tilde{\boldsymbol{\theta}}} < 0$, $\mathcal{L}(\boldsymbol{\beta}, \sigma^2; \mathbf{y})$ ($\mathcal{L}(\boldsymbol{\beta}, \sigma^2; \mathbf{y})$) tiene un máximo en $\tilde{\boldsymbol{\theta}} = \begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\sigma}^2 \end{pmatrix}$.

Note que sólo es necesario encontrar a $\tilde{\boldsymbol{\beta}}$, pues una vez hecho esto, $\tilde{\sigma}^2$ se obtiene de inmediato.

Resumiendo:

TEOREMA C.1. *Los estimadores de verosimilitud máxima de $\boldsymbol{\beta}$ y σ^2 son, respectivamente:*

i.

$$\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} = \mathbf{X}^- \mathbf{y},$$

ii. y

$$\begin{aligned} \tilde{\sigma}^2 &= \frac{1}{n} \|\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}\|^2 = \frac{1}{n} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})' (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}) = \frac{1}{n} \mathbf{y}' (\mathbf{I} - \mathbf{X}\mathbf{X}^-) \mathbf{y} = \frac{1}{n} (\mathbf{y}'\mathbf{X} - \tilde{\boldsymbol{\beta}}\mathbf{X}'\mathbf{y}) \\ &= \frac{1}{n} \|\mathbf{I} - \mathbf{X}\mathbf{X}^-\|^2. \end{aligned}$$

Algunas de las propiedades de los estimadores máximo verosímiles son:

TEOREMA C.2.

i.

$$\tilde{\boldsymbol{\beta}} \sim N_p(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}).$$

ii.

$$\frac{n\tilde{\sigma}^2}{\sigma^2} \sim \chi_{n-p}^2.$$

iii.

$\tilde{\boldsymbol{\beta}}$ y $\tilde{\sigma}^2$ son independientes (ver Apéndice A).

Más aún, note que

$E(\tilde{\boldsymbol{\beta}}) = \boldsymbol{\beta}$ y definiendo $\hat{\sigma}^2 = \frac{n\tilde{\sigma}^2}{n-p}$ se tiene que

$$E(\hat{\sigma}^2) = \frac{E(\mathbf{y}'(\mathbf{I} - \mathbf{X}\mathbf{X}^{-})\mathbf{y})}{n-p} = \frac{\sigma^2(n-p)}{n-p} = \sigma^2.$$

Escribiendo $\tilde{\boldsymbol{\theta}} = \begin{pmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\sigma}^2 \end{pmatrix}$, $\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\sigma}^2 \end{pmatrix}$ es un estimador insesgado de $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\beta} \\ \sigma^2 \end{pmatrix}$.

Estimación de formas lineales.

Considere el modelo lineal general de rango completo

$$\mathbf{y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}_n) \text{ (ver Apéndice B).}$$

Es de interés estimar la forma lineal $\boldsymbol{\ell}'\boldsymbol{\beta}$, donde $\boldsymbol{\ell} \in \mathfrak{R}^p$ es un vector constante. El estimador natural es $\boldsymbol{\ell}'\hat{\boldsymbol{\beta}}$ y

TEOREMA C.3. Si $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\beta} \\ \sigma^2 \end{pmatrix}$, entonces para todo $\boldsymbol{\theta}$,

- i. $E_{\theta}(\ell' \hat{\beta}) = \ell' \beta.$
- ii. $\text{Cov}_{\theta}(\ell' \hat{\beta}) = \sigma^2 \ell' (X' X)^{-1} \ell.$
- iii. $\ell' \hat{\beta} \sim N(\ell' \beta, \sigma^2 \ell' (X' X)^{-1} \ell).$

Como consecuencia, $\ell' \hat{\beta}$ es el estimador máximo verosímil de $\ell' \beta$.

Adicionalmente note que dado $\theta = \begin{pmatrix} \beta \\ \sigma^2 \end{pmatrix}$, el estimador $\hat{\theta} = \begin{pmatrix} \hat{\beta} \\ \hat{\sigma}^2 \end{pmatrix}$, además de ser insesgado, tiene las siguientes propiedades:

- 1) Suficiencia. El estimador $\hat{\theta}$ es suficiente, esto es, si una persona dispone de $\hat{\theta}$ y no dispone del vector y , puede hacer tan buenas inferencias como una persona que dispone del vector y .
- 2) Completez. Si se tienen dos estimadores insesgados de una función paramétrica $G(\theta)$ y ambos estimadores son funciones de $\hat{\theta} = \begin{pmatrix} \hat{\beta} \\ \hat{\sigma}^2 \end{pmatrix}$, entonces los dos estimadores coinciden. Esto es, si $T_1(\cdot)$ y $T_2(\cdot)$ son los estimadores insesgados de $G(\theta)$

$$E_{\theta} \left(T_1(\hat{\beta}, \hat{\sigma}^2) \right) = E_{\theta} \left(T_2(\hat{\beta}, \hat{\sigma}^2) \right)$$

Para todo θ , entonces

$$T_1(\hat{\beta}, \hat{\sigma}^2) = T_2(\hat{\beta}, \hat{\sigma}^2)$$

Prueba de una hipótesis lineal.

Ahora veremos cómo efectuar pruebas de hipótesis acerca del parámetro β . La única restricción es que tal hipótesis se exprese mediante condiciones lineales sobre el parámetro β . En forma precisa, las hipótesis a probar se expresan como

$$H\beta = h,$$

Donde $H \in \mathfrak{R}^{q \times p}$ y $h \in \mathfrak{R}^q$ los cuales son fijos y conocidos.

Ejemplo. Suponga que

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i, \quad i = 1, 2, \dots, n$$

Se desean probar las siguientes hipótesis

- i. $\beta_0 = 0$
- ii. $\beta_0 = \beta_1 = \beta_2$
- iii. $\beta_1 + \beta_2 = 2\beta_3$
- iv. $\beta_1 + 2\beta_2 = 3\beta_3$
- v. $\beta_2 = 5\beta_3$
- vi. $\beta_0 = 5$

En notación matricial estas hipótesis se expresan como:

i.

$$[1 \ 0 \ 0 \ 0] \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = 0$$

$$H \quad \beta = h$$

ii.

$$\begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{o} \quad \begin{bmatrix} 0 & 0 & -1 & 1 \\ 0 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

iii.

$$\begin{bmatrix} 0 & 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = 0$$

iv.

$$\begin{bmatrix} 0 & 1 & 2 & -3 \\ 0 & 0 & 1 & -5 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix}$$

En adelante se consideran hipótesis de la forma

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{h}$$

Donde

i.

$\mathbf{H}\boldsymbol{\beta} = \mathbf{h}$, es un sistema de ecuaciones lineales consistente.

ii.

$\mathbf{H} \in \mathfrak{R}^{q \times p}$, con $r(\mathbf{H}) = q \leq p$, y $\mathbf{h} \in \mathfrak{R}^{q \times 1}$.

Una vez observado el vector \mathbf{y} , se desea tomar una decisión sobre aceptar o rechazar la hipótesis. Lo más intuitivo es rechazar la hipótesis $\mathbf{H}\boldsymbol{\beta} = \mathbf{h}$ si $\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{h}$ es grande. Desde luego un vector es grande si su norma lo es.

Recordando que

$$\hat{\boldsymbol{\beta}} \sim N_p(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}).$$

Se tiene

$$\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{h} \sim N_q(\mathbf{H}\boldsymbol{\beta} - \mathbf{h}, \sigma^2\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}').$$

Entonces

$$\frac{(\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{h})' [\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}']^{-1} (\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{h})}{\sigma^2} \sim \chi_{q,\lambda}^2.$$

Donde

$$\lambda = \frac{1}{2\sigma^2} (\mathbf{H}\boldsymbol{\beta} - \mathbf{h})' [\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}']^{-1} (\mathbf{H}\boldsymbol{\beta} - \mathbf{h}).$$

Además, recuerde que

$$\frac{(n-p)\hat{\sigma}^2}{\sigma^2} \sim \chi_{n-p}^2.$$

Entonces

$$W = \frac{(\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{h})' [\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}']^{-1} (\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{h})}{\sigma^2 q} \frac{(n-p)\hat{\sigma}^2}{(n-p)\sigma^2}$$

$$= \frac{(\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{h})' [\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}]^{-1} (\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{h})}{q\hat{\sigma}^2} \sim F_{q,n-p,\lambda}.$$

Ahora observe que cuando $\mathbf{H}\boldsymbol{\beta} = \mathbf{h}$, $\lambda = 0$ y entonces $W \sim F_{q,n-p}$ (centrada) como lo muestra la Figura 3.

Así, se tiene la siguiente regla de decisión: Rechace la hipótesis $\mathbf{H}\boldsymbol{\beta} = \mathbf{h}$ cuando

$$W > F_{q,n-q}(\alpha).$$

Esto produce una prueba con nivel de significación α .

Invariabilidad de la prueba de hipótesis $\mathbf{H}\boldsymbol{\beta} = \mathbf{h}$.

Para probar la hipótesis $\mathbf{H}\boldsymbol{\beta} = \mathbf{h}$ a un nivel de significación α dado, se procede como sigue:

Observación: recuerde que el nivel de significación de una prueba de hipótesis, es la probabilidad de rechazar la hipótesis cuando ésta es cierta.

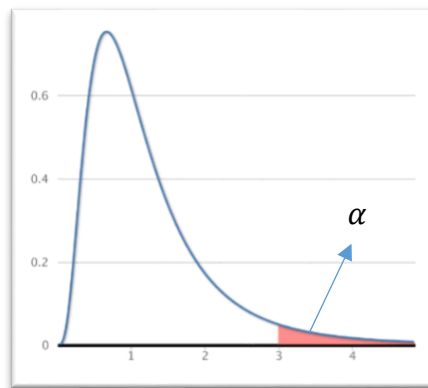
Calcule

$$W = \frac{(\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{h})' [\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}]^{-1} (\mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{h})}{q\hat{\sigma}^2}.$$

Bajo la condición de que $\mathbf{H}\boldsymbol{\beta} = \mathbf{h}$,

$$W \sim F_{q,n-p}.$$

Se acepta la hipótesis $\mathbf{H}\boldsymbol{\beta} = \mathbf{h}$ si $W \leq F_{q,n-p}(\alpha)$.



$F_{q,n-q}(\alpha)$

Figura 3. Distribución $F_{q,n-q}(\alpha)$ para probar $H\beta = h$

Se rechaza la hipótesis $\mathbf{H}\boldsymbol{\beta} = \mathbf{h}$ si $W > F_{q,n-p}(\alpha)$, en este caso se dice que se acepta $\mathbf{H}\boldsymbol{\beta} \neq \mathbf{h}$. Así, este procedimiento produce una regla de decisión que selecciona cuál de las hipótesis $\mathbf{H}\boldsymbol{\beta} = \mathbf{h}$ o $\mathbf{H}\boldsymbol{\beta} \neq \mathbf{h}$ debe aceptarse.

En este contexto, $\mathbf{H}\boldsymbol{\beta} = \mathbf{h}$ se llama hipótesis nula (pues establece que $\mathbf{H}\boldsymbol{\beta} - \mathbf{h} = \mathbf{0}$ vector nulo) y se denota por H_0 . Mientras que la hipótesis $\mathbf{H}\boldsymbol{\beta} \neq \mathbf{h}$ se llama hipótesis alternativa y se denota por H_a .

La regla de decisión anterior se refiere como una prueba de

$$H_0: \mathbf{H}\boldsymbol{\beta} = \mathbf{h} \quad \text{versus} \quad H_a: \mathbf{H}\boldsymbol{\beta} \neq \mathbf{h}.$$

Si $\mathbf{H}\boldsymbol{\beta} = \mathbf{h}$ y $\mathbf{H}_1\boldsymbol{\beta} = \mathbf{h}_1$ son dos sistemas de ecuaciones que establecen las mismas condiciones sobre $\boldsymbol{\beta}$, entonces los dos sistemas son equivalentes, por lo tanto existe una matriz $\mathbf{Q} \in \mathfrak{R}^{q \times q}$ invertible, tal que

$$\mathbf{QH} = \mathbf{H}_1 \text{ y } \mathbf{Qh} = \mathbf{h}_1. \quad (*)$$

TEOREMA C.4. *Suponga que \mathbf{H} y \mathbf{H}_1 son como antes si W y W_1 son estadísticos obtenidos usando (\mathbf{H}, \mathbf{h}) y $(\mathbf{H}_1, \mathbf{h}_1)$ respectivamente, entonces $W = W_1$.*

Demostración. Veamos que existe \mathbf{Q} tal que (*) se cumple, de

$$\mathbf{QH} = \mathbf{H}_1$$

$$\Rightarrow \mathbf{QHH}' = \mathbf{H}_1\mathbf{H}'.$$

Ahora note que $\mathbf{H} \in \mathfrak{R}^{q \times p}$, $r(\mathbf{H}) = r(\mathbf{HH}') = q$, luego $\mathbf{HH}' > \mathbf{0}$, así

$$\begin{aligned} \mathbf{Q} &= \mathbf{H}_1\mathbf{H}'(\mathbf{HH}')^{-1} = \mathbf{H}_1\mathbf{H}'(\mathbf{HH}')^{-} = \mathbf{H}_1\mathbf{H}'\mathbf{H}'^{-}\mathbf{H}^{-} = \mathbf{H}_1(\mathbf{H}^{-}\mathbf{H})'\mathbf{H}^{-} = \mathbf{H}_1\mathbf{H}^{-}\mathbf{HH}^{-} \\ &= \mathbf{H}_1\mathbf{H}^{-}. \end{aligned}$$

Además, note que como $H\beta = h$, entonces

$$Qh = H_1 H^{-1} h = H_1 \beta = h_1.$$

Ahora

$$W_1 = \frac{(H_1 \hat{\beta} - h_1)' (H_1 (X'X)^{-1} H_1')^{-1} (H_1 \hat{\beta} - h_1)'}{q \hat{\sigma}^2}.$$

y

$$\begin{aligned} W &= \frac{(Q(H\hat{\beta} - h))' [QH(X'X)^{-1}H'Q]^{-1} (QH\hat{\beta} - h)}{q \hat{\sigma}^2}, \\ &= \frac{(H\hat{\beta} - h)' Q' Q^{-1} (H_1 (X'X)^{-1} H_1')^{-1} Q^{-1} Q (H\hat{\beta} - h)}{q \hat{\sigma}^2}, \\ &= \frac{(H\hat{\beta} - h)' I (H_1 (X'X)^{-1} H_1')^{-1} I (H\hat{\beta} - h)}{q \hat{\sigma}^2}, \\ &= \frac{(H\hat{\beta} - h)' (H_1 (X'X)^{-1} H_1')^{-1} (H\hat{\beta} - h)}{q \hat{\sigma}^2}. \end{aligned}$$

Intervalos de confianza y pruebas de hipótesis.

Sea $\ell \in \mathbb{R}^p$ se puede construir un intervalo de confianza para $\ell' \beta$ como sigue

$$\ell' \hat{\beta} \sim N_1(\ell' \beta, \sigma^2 \ell' (X'X)^{-1} \ell).$$

Luego

$$\frac{\ell' \hat{\beta} - \ell' \beta}{\sigma \sqrt{\ell' (X'X)^{-1} \ell}} \sim N(0,1),$$

y

$$\frac{(n-p)\hat{\sigma}^2}{\sigma^2} \sim \chi_{n-p}^2.$$

Así

$$T = \frac{\frac{\boldsymbol{\ell}'\hat{\boldsymbol{\beta}} - \boldsymbol{\ell}'\boldsymbol{\beta}}{\sigma\sqrt{\boldsymbol{\ell}'(\mathbf{X}\mathbf{X})^{-1}\boldsymbol{\ell}}}}{\sqrt{\frac{(n-p)\hat{\sigma}^2}{\sigma^2(n-p)}}} = \frac{\boldsymbol{\ell}'\hat{\boldsymbol{\beta}} - \boldsymbol{\ell}'\boldsymbol{\beta}}{\hat{\sigma}\sqrt{\boldsymbol{\ell}'(\mathbf{X}\mathbf{X})^{-1}\boldsymbol{\ell}}} = \frac{\boldsymbol{\ell}'\hat{\boldsymbol{\beta}} - \boldsymbol{\ell}'\boldsymbol{\beta}}{\sqrt{\widehat{\text{var}}(\boldsymbol{\ell}'\hat{\boldsymbol{\beta}})}} \sim t_{n-p}.$$

Sigue una distribución t para todo θ , así

$$P_{\theta} = [|T| \leq t_{n-p,(\alpha/2)}] = 1 - \alpha$$

$$\Rightarrow P_{\theta} = \left[\left| \frac{\boldsymbol{\ell}'\hat{\boldsymbol{\beta}} - \boldsymbol{\ell}'\boldsymbol{\beta}}{\sqrt{\widehat{\text{var}}(\boldsymbol{\ell}'\hat{\boldsymbol{\beta}})}} \right| \leq t_{n-p,(\alpha/2)} \right] = 1 - \alpha$$

$$\Rightarrow P_{\theta} = \left[\boldsymbol{\ell}'\hat{\boldsymbol{\beta}} - \sqrt{\widehat{\text{var}}(\boldsymbol{\ell}'\hat{\boldsymbol{\beta}})}t_{n-p,(\alpha/2)} \leq \boldsymbol{\ell}'\boldsymbol{\beta} \leq \boldsymbol{\ell}'\hat{\boldsymbol{\beta}} + \sqrt{\widehat{\text{var}}(\boldsymbol{\ell}'\hat{\boldsymbol{\beta}})}t_{n-p,(\alpha/2)} \right] = 1 - \alpha.$$

Concluyendo que el intervalo aleatorio

$$\left(\boldsymbol{\ell}'\hat{\boldsymbol{\beta}} - \sqrt{\widehat{\text{var}}(\boldsymbol{\ell}'\hat{\boldsymbol{\beta}})}t_{n-p,(\alpha/2)}, \boldsymbol{\ell}'\hat{\boldsymbol{\beta}} + \sqrt{\widehat{\text{var}}(\boldsymbol{\ell}'\hat{\boldsymbol{\beta}})}t_{n-p,(\alpha/2)} \right).$$

Es un intervalo de confianza para $\boldsymbol{\ell}'\boldsymbol{\beta}$, con un nivel de confianza de $1 - \alpha$.

Observación.

Note que $\widehat{\text{var}}(\boldsymbol{\ell}'\hat{\boldsymbol{\beta}}) = \hat{\sigma}^2\boldsymbol{\ell}'(\mathbf{X}\mathbf{X})^{-1}\boldsymbol{\ell}$.

Ahora considere el problema de probar la hipótesis

$$H_0: \boldsymbol{\ell}'\boldsymbol{\beta} = \ell_0 \quad \text{vs} \quad H_a: \boldsymbol{\ell}'\boldsymbol{\beta} = \ell_0.$$

Procediendo como en la sección anterior se tiene:

1. Calcular

$$W = \frac{(\ell' \hat{\beta} - \ell_0)' (\ell' (X'X)^{-1} \ell)^{-1} (\ell' \hat{\beta} - \ell_0)'}{1 \hat{\sigma}^2}$$

$$= \frac{(\ell' \hat{\beta} - \ell_0)^2}{\hat{\sigma}^2 (\ell' (X'X)^{-1} \ell)} = \frac{(\ell' \hat{\beta} - \ell_0)^2}{\widehat{\text{var}}(\ell' \hat{\beta})}$$

2. Se cumple H_0 si y sólo si,

3.

$$W \leq F_{1,n-p,\alpha}$$

Pero recuerde que $F_{1,n-p,\alpha} = t_{n-p,(\alpha/2)}^2$, entonces $W \leq F_{1,n-p,\alpha}$ es equivalente

a

$$\frac{(\ell' \hat{\beta} - \ell_0)^2}{\widehat{\text{var}}(\ell' \hat{\beta})} \leq t_{n-p,(\alpha/2)}^2 \quad \text{ó}$$

$$\left| \frac{\ell' \hat{\beta} - \ell_0}{\sqrt{\widehat{\text{var}}(\ell' \hat{\beta})}} \right| \leq t_{n-p,(\alpha/2)} \quad \text{ó}$$

$$\ell' \hat{\beta} - \sqrt{\widehat{\text{var}}(\ell' \hat{\beta})} t_{n-p,(\alpha/2)} \leq \ell_0 \leq \ell' \hat{\beta} + \sqrt{\widehat{\text{var}}(\ell' \hat{\beta})} t_{n-p,(\alpha/2)}$$

Así,

H_0 se acepta con un nivel de significancia α , si y sólo si ℓ_0 pertenece al intervalo de confianza de $100(1 - \alpha)\%$ para $\ell' \beta$.

Prueba de la hipótesis $H\boldsymbol{\beta} = \mathbf{h}$ usando el método de la razón de verosimilitud

Recuerde que

$$\Omega = \{(\boldsymbol{\beta}, \sigma^2) | \boldsymbol{\beta} \in \mathfrak{R}^p, \sigma^2 > 0\}.$$

Y denote

$$\omega = \{(\boldsymbol{\beta}, \sigma^2) | \sigma^2 > 0, H\boldsymbol{\beta} = \mathbf{h}\}.$$

Denotando el espacio paramétrico reducido por la hipótesis ahora considere

$$\hat{L}_\omega(\mathbf{y}) = \max L(\mathbf{y}, \hat{\boldsymbol{\theta}}_\omega) \text{ y } \hat{L}_\Omega(\mathbf{y}) = \max L(\mathbf{y}, \hat{\boldsymbol{\theta}}_\Omega).$$

Donde $\boldsymbol{\theta} \in \omega$ y $\boldsymbol{\theta} \in \Omega$ respectivamente.

La razón de verosimilitud está dada por:

$$\lambda(\mathbf{y}) = \frac{\hat{L}_\omega(\mathbf{y})}{\hat{L}_\Omega(\mathbf{y})}.$$

La siguiente prueba de $H_0: H\boldsymbol{\beta} = \mathbf{h}$ o $H_0: \boldsymbol{\theta} \in \omega$ es llamada prueba de la razón de verosimilitud.

Se rechaza $H_0: \boldsymbol{\theta} \in \omega$ ($H_0: H\boldsymbol{\beta} = \mathbf{h}$) si $\lambda(\mathbf{y}) < c$ dado un nivel de significación α .

En adelante considere que $H \in \mathfrak{R}^{q \times p}$, $r(H) = q$, notando que

$$\hat{L}_\Omega(\mathbf{y}) = \left[\exp\left(-\frac{1}{2} \|\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_\Omega\|^2 / \hat{\sigma}_\Omega^2\right) \right] \frac{1}{(2\pi\hat{\sigma}_\Omega^2)^{n/2}} \exp(-n/2),$$

$$= \frac{1}{(2\pi\hat{\sigma}_\Omega^2)^{n/2}} \exp(-n/2).$$

Donde $\hat{\boldsymbol{\beta}}_\Omega = \mathbf{X}^{-1}\mathbf{y}$ y $\hat{\sigma}_\Omega^2 = \frac{\|\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_\Omega\|^2}{n}$.

$$\hat{L}_\omega(\mathbf{y}) = \max L(\mathbf{y}, \boldsymbol{\theta}), \quad \boldsymbol{\theta} \in \omega = \{(\boldsymbol{\beta}, \sigma^2) | \sigma^2 > 0, \mathbf{H}\boldsymbol{\beta} = \mathbf{h}\},$$

Donde $\boldsymbol{\theta} \in \omega$.

Esto es equivalente a

$$\max L(\mathbf{y}, \boldsymbol{\theta})$$

Sujeto a

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{h}.$$

Usando multiplicadores de Lagrange se tiene:

$$\begin{aligned} L &= L(\mathbf{y}, \boldsymbol{\theta}) - \boldsymbol{\lambda}'(\mathbf{H}\boldsymbol{\beta} - \mathbf{h}) \\ &= (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2\right) - \boldsymbol{\lambda}'(\mathbf{H}\boldsymbol{\beta} - \mathbf{h}). \end{aligned}$$

Derivando se obtiene

1.

$$\frac{\partial L}{\partial \boldsymbol{\beta}} = \frac{1}{2\sigma^2} (-2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta})L(\mathbf{y}, \boldsymbol{\theta}) - \mathbf{H}'\boldsymbol{\lambda}.$$

2.

$$\frac{\partial L}{\partial \sigma^2} = -\frac{n}{2}(\sigma^2)^{-\frac{n}{2}-1}(2\pi)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2\right)$$

$$\begin{aligned}
& + (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2\right) \left[\frac{1}{2\sigma^4} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2\right], \\
& = \left[-\frac{n}{2}(\sigma^2)^{-1} + \frac{1}{2\sigma^4} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2\right] L(\mathbf{y}, \boldsymbol{\theta}).
\end{aligned}$$

3.

$$\frac{\partial L}{\partial \boldsymbol{\lambda}} = -(\mathbf{H}\boldsymbol{\beta} - \mathbf{h}).$$

Sean $\tilde{\boldsymbol{\beta}}_\omega, \tilde{\sigma}_\omega^2$ y $\tilde{\boldsymbol{\lambda}}$ soluciones al sistema anterior luego de la derivada 1, se tiene:

$$\begin{aligned}
& \frac{1}{\tilde{\sigma}_\omega^2} [\mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\tilde{\boldsymbol{\beta}}_\omega] \hat{L}(\mathbf{y}, \tilde{\boldsymbol{\theta}}_\omega) - \mathbf{H}'\tilde{\boldsymbol{\lambda}} = 0 \\
& \Rightarrow \mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\tilde{\boldsymbol{\beta}}_\omega = \frac{\tilde{\sigma}_\omega^2 \mathbf{H}'\tilde{\boldsymbol{\lambda}}}{\hat{L}(\mathbf{y}, \tilde{\boldsymbol{\theta}}_\omega)}.
\end{aligned}$$

Defina

$$\boldsymbol{\lambda}^* = \frac{\tilde{\sigma}_\omega^2 \tilde{\boldsymbol{\lambda}}}{\hat{L}(\mathbf{y}, \tilde{\boldsymbol{\theta}}_\omega)}.$$

Por lo tanto

$$\mathbf{H}'\boldsymbol{\lambda}^* + \mathbf{X}'\mathbf{X}\tilde{\boldsymbol{\beta}}_\omega = \mathbf{X}'\mathbf{y}.$$

De la ecuación 3. $\mathbf{H}\tilde{\boldsymbol{\beta}}_\omega - \mathbf{h} = \mathbf{0}$ luego, estas dos condiciones (ecuaciones) se pueden escribir como:

4.

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{H}' \\ \mathbf{H} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{\beta}}_\omega \\ \boldsymbol{\lambda}^* \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{h} \end{bmatrix}.$$

Note que $L(\mathbf{y}, \tilde{\boldsymbol{\theta}}_\omega) > 0$ luego por 2

$$-\frac{n}{2}(\tilde{\sigma}_\omega^2)^{-1} + \frac{1}{2}(\tilde{\sigma}_\omega^4)^{-1}\|\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_\omega\|^2 = 0$$

$$\Rightarrow \|\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_\omega\|^2 = n(\tilde{\sigma}_\omega^2)^{-1}\tilde{\sigma}_\omega^4$$

$$\Rightarrow \tilde{\sigma}_\omega^2 = \frac{1}{n}\|\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}_\omega\|^2.$$

Para resolver el sistema 4, recuerde que dada \mathbf{A} no singular

$$\begin{aligned} \mathbf{A}^{-1} &= \begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{H}' \\ \mathbf{H} & \mathbf{0} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} (\mathbf{X}'\mathbf{X})^{-1} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}'[-\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}']^{-1}\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1} & -(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}'(-\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}')^{-1} \\ (-\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}')^{-1}\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1} & (-\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}')^{-1} \end{pmatrix}. \end{aligned}$$

Por lo tanto

$$\begin{pmatrix} \tilde{\boldsymbol{\beta}}_\omega \\ \lambda^* \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{H}' \\ \mathbf{H} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{h} \end{pmatrix}.$$

Así

$$\tilde{\boldsymbol{\beta}}_\omega = \mathbf{X}^{-}\mathbf{y} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}'[\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}']^{-1}\mathbf{H}\mathbf{X}^{-}\mathbf{y} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}'(\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}')^{-1}\mathbf{h}$$

$$\lambda^* = (\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}')^{-1}\mathbf{H}\mathbf{X}^{-}\mathbf{y} - (\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}')^{-1}\mathbf{h}.$$

Si $\tilde{\boldsymbol{\beta}}_\Omega = \mathbf{X}^{-}\mathbf{y}$ se tiene

$$\lambda^* = [\mathbf{H}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{H}']^{-1}(\mathbf{H}\tilde{\boldsymbol{\beta}}_\Omega - \mathbf{h})$$

y

$$\tilde{\beta}_\omega = \tilde{\beta}_\Omega - (X'X)^{-1}H'\lambda^*.$$

Ahora

$$\tilde{\sigma}_\omega^2 = \frac{1}{n} \|y - X\tilde{\beta}_\omega\|^2 = \frac{1}{n} (y'y - 2y'X\tilde{\beta}_\omega + \tilde{\beta}_\omega'X'X\tilde{\beta}_\omega).$$

Observe que

$$y'X\tilde{\beta}_\omega = y'X[\tilde{\beta}_\Omega - (X'X)^{-1}H'\lambda^*] = y'X\tilde{\beta}_\Omega - y'X(X'X)^{-1}H'\lambda^* = y'X\tilde{\beta}_\Omega - \tilde{\beta}'_\Omega H'\lambda^*, y$$

$$\begin{aligned} \tilde{\beta}_\omega'X'X\tilde{\beta}_\omega &= [\lambda^*H(X'X)^{-1} - \tilde{\beta}'_\Omega]X'X[(X'X)^{-1}H'\lambda^* - \tilde{\beta}_\Omega] \\ &= \lambda^*H(X'X)^{-1}H'\lambda^* + \tilde{\beta}'_\Omega(X'X)\tilde{\beta}_\Omega - 2\tilde{\beta}'_\Omega H'\lambda^*. \end{aligned}$$

Por lo tanto

$$\begin{aligned} \tilde{\sigma}_\omega^2 &= \frac{1}{n} [y'y - 2[y'X\tilde{\beta}_\Omega - \tilde{\beta}'_\Omega H'\lambda^*] + \lambda^*H(X'X)^{-1}H'\lambda^* + \tilde{\beta}'_\Omega(X'X)\tilde{\beta}_\Omega - 2\tilde{\beta}'_\Omega H'\lambda^*] \\ &= \tilde{\sigma}_\Omega^2 + \lambda^*H(X'X)^{-1}H'\lambda^* + 2\tilde{\beta}'_\Omega H'\lambda^* - 2\tilde{\beta}'_\Omega H'\lambda^* \\ &= \tilde{\sigma}_\Omega^2 + \lambda^*H(X'X)^{-1}H'\lambda^* + 0 + 0 \\ &= \tilde{\sigma}_\Omega^2 + \frac{1}{n} \lambda^*H(X'X)^{-1}H'\lambda^* \\ &= \tilde{\sigma}_\Omega^2 + \frac{1}{n} [H\tilde{\beta}_\Omega - h]' [H(X'X)^{-1}H']^{-1} H(X'X)^{-1}H [H(X'X)^{-1}H']^{-1} [H\tilde{\beta}_\Omega - h] \\ &= \tilde{\sigma}_\Omega^2 + [H\tilde{\beta}_\Omega - h]' (H(X'X)^{-1}H')^{-1} [H\tilde{\beta}_\Omega - h] \frac{1}{n}. \end{aligned}$$

Usando la razón de verosimilitud generalizada, una función de $\lambda(y)$,

$$W = \binom{n-p}{q} \left(\frac{\tilde{\sigma}_\omega^2 - \tilde{\sigma}_\Omega^2}{\sigma_\Omega^2} \right) = \frac{\frac{1}{n} [H\hat{\beta} - h]' (H(X'X)^{-1}H')^{-1} [H\hat{\beta} - h]}{\tilde{\sigma}_\Omega^2} \binom{n-p}{q}$$

Donde $\hat{\beta} = X^{-1}y$ y $\hat{\sigma}_{\Omega}^2 = \frac{1}{n} \|y - X\hat{\beta}\|^2$

$$\begin{aligned} W &= \frac{\frac{1}{n} [H\hat{\beta} - h]' (H(X'X)^{-1}H')^{-1} [H\hat{\beta} - h]}{\frac{\|y - X\hat{\beta}\|^2}{n}} \left(\frac{n-p}{q}\right), \\ &= \frac{[H\hat{\beta} - h]' (H(X'X)^{-1}H')^{-1} [H\hat{\beta} - h]}{\|y - X\hat{\beta}\|^2} \left(\frac{n-p}{q}\right), \\ &= \frac{[H\hat{\beta} - h]' (H(X'X)^{-1}H')^{-1} [H\hat{\beta} - h]}{q \frac{\|y - X\hat{\beta}\|^2}{n-p}}. \end{aligned}$$

De donde finalmente

$$W = \frac{[H\hat{\beta} - h]' (H(X'X)^{-1}H')^{-1} [H\hat{\beta} - h]}{q\hat{\sigma}^2},$$

con $(n-p)\hat{\sigma}^2 = \|y - X\hat{\beta}\|^2$, $\hat{\beta} = X^{-1}y$.

Resumiendo

TEOREMA C.5. *Bajo el modelo lineal general de rango completo y dada la hipótesis general:*

$$H_0: H\beta = h \quad \text{versus} \quad H_a: H\beta \neq h,$$

Se rechaza H_0 si $W \geq F_{\alpha, n-p, q}$. Donde

$$W = \frac{(H\hat{\beta} - h)' (H(X'X)^{-1}H')^{-1} [H\hat{\beta} - h]}{q\hat{\sigma}^2}, \quad \hat{\beta} = X^{-1}y, \quad \hat{\sigma}^2 = \|y - X\hat{\beta}\|^2 / n - p.$$

APÉNDICE D

Secuencia para comparación de modelos en el software R

```
#####
```

```
#           Comparación de modelos
```

```
#           Tesina Arwell
```

```
#           2018
```

```
#####
```

```
rm(list=ls(all=TRUE)) # Limpia la memoria
```

```
# library(MASS) sólo incluirla una vez
```

```
library(MASS)
```

```
#####
```

```
#           Datos
```

```
#####
```

```
Y1 <- c(0.1,0,0.1,3.4,18.0,4.7,15,43,71.9,21.4,2.1,0.2)
```

```
Y1
```

```
Y2 <- c(0.2,0.1,0.1,1.0,7.9,7.0,15.8,41.2,75.7,6.4,3.1,0)
```

```
Y2
```

```
Y3 <- c(0,0,0,0,0,0,0,0.6,1.3,0.1,0.2,0)
```

```
Y3
```

```
T <- c(13.7,14.7,16.5,19,21.3,21,19.6,19.3,18.8,17.2,15.9,14.4)
```

```
T
```

```
P <- c(17,5.7,4.5,9.5,27.2,129.3,162.9,168,106.7,30.3,9.7,5.6)
```

```
P
```

```
Y <- cbind(Y1,Y2,Y3)
```

```
Y
```

```
X <- cbind(rep(1,12),T,P,T*T,P*P,T*P,T*P*P)
```

```
X
```

```
#####
```

```
#      Análisis de regresión múltiple
```

```
#      A través de análisis de regresión multivariada
```

```
#####
```

```
MMReg=manova(Y~X)
```

```
MMReg
```

```
ct=c("Pillai", "Wilks", "Hotelling-Lawley", "Roy")
```

```
ct
```

```
MMReg1=summary(MMReg,test=ct[4])
```

```
MMReg1
```

```
SSE=MMReg1$SS$Residuals
```

```
SSE
```

```
SSReg=MMReg1$SS$X
```

```
SSReg
```

```
SSTotal=SSReg+SSE
```

```
SSTotal
```

```
Roy=eigen(ginv(SSE)%*%SSReg)
```

```
Roy
```

```
# Anovas
```

```
Anovas=summary.aov(MMReg)
```

```
Anovas
```

```
# Coeficientes del modelo
```

```
BB=as.matrix((coefficients(MMReg)),7)
```

```
BB
```

```
# Valores predichos
```

```
pred=fitted(MMReg)
```

```
pred
```

```
# Residuales
```

```

res=residuals(MMReg)

res

#####

#   Medidas de asociación

#####

# Coeficientes de determinación

R2=diag(SSReg/SSTotal)

R2

# Coeficientes de correlación

r=sqrt(R2)

r

#####

#   Prueba para la igualdad de modelos

#   Ver Graybill (1976), Sección 8.6.2

#   Modelo completo, i.e., incluyendo  $b_{\{0\}}$ 

#####

H=3 # Número de modelos a comparar

N=36 # 12x3, Número de observaciones por el número de modelos

```

p=7 # Número de parámetros en el modelo

X1=cbind(rep(1,12),T,P,T*T,P*P,T*P,T*P*P)

X1

XtX=t(X1)%*%X1

XtX

XGX=X1%*%ginv(X1)

XGX

suma1=0

for(h in 1:H)

 suma1=suma1 + t(Y[,h])%*%XGX%*%Y[,h]

suma1

suma2=rep(0,7)

for(h in 1:H)

 suma2=suma2 + t(X1)%*%Y[,h]

suma2

suma3=t(suma2)%*%ginv(H*XtX)%*%suma2

suma3

suma4=0

```

for(h in 1:H)

  suma4=suma4 + t(Y[,h])%*%Y[,h]

suma4

prod1=suma1-suma3/(suma4-suma1)

prod1

prod2=(N-H*p)/((H-1)*p)

prod2

W=prod1*prod2 #

W

F05=qf(0.95,(H-1)*p,N-H*p)

F05

F01=qf(0.99,(H-1)*p,N-H*p)

F01

# pvalue = 0

pvalue=1-pf(W,(H-1)*p,N-H*p)

pvalue

### Los modelos son diferentes

#####

```

```
# Optimización de la superficie de respuesta
```

```
# Max Y1
```

```
# Max Y2
```

```
# Max Y3
```

```
#####
```

```
# Note que la función para B[,1] debe ser maximizada,
```

```
# Por lo tanto, "supresp" es definida como la función (x,b){-()}, en otro caso
```

```
# (minimización), "supresp" la función es(x,b){}
```

```
#####
```

```
lix1=min(T)
```

```
lix1
```

```
lsx1=max(T)
```

```
lsx1
```

```
lix2=min(P)
```

```
lix2
```

```
lsx2=max(P)
```

```
lsx2
```

```
IOP=rep(0,3) # Individual optima
```


IOP

```
supresp=function(x,b){-
```

```
(b[1]+b[2]*x[1]+b[3]*x[2]+b[4]*x[1]^2+b[5]*x[2]^2+b[6]*x[1]*x[2]+b[7]*x[1]*x[2]*x[2])
```

```
# Max
```

```
supresp
```

```
#####
```

```
# Maximizando la variable Y1
```

```
#####
```

```
ORSY1=nlminb(c(0,0), supresp,b =BB[,1], lower = c(lix1,lix2), upper = c(lsx1,lsx2))
```

```
ORSY1
```

```
CPY1=ORSY1$par
```

```
CPY1
```

```
IOP[1]=-ORSY1$objective
```

```
IOP[1]
```

```
#####
```

```
# Maximizando la variable Y2
```

```
#####
```

```
ORSY2=nlminb(c(0, 0), supresp, b = BB[,2], lower = c(lix1,lix2), upper = c(lsx1,lsx2))
```

```
ORSY2
```

CPY2=ORSY2\$par

CPY2

IOP[2]=-ORSY2\$objective

IOP[2]

#####

Maximizando la variable Y3

#####

ORSY3=nlminb(c(0, 0), supresp, b = BB[,3], lower = c(lix1,lix2), upper = c(lsx1,lsx2))

ORSY3

CPY3=ORSY3\$par

CPY3

IOP[3]=-ORSY3\$objective

IOP[3]

