VNiVERSiDAD D SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL

GRADO EN ESTADÍSTICA

FACULTAD DE CIENCIAS

CRITERIOS DE OPTIMIZACIÓN CARACTERÍSTICOS

Trabajo de fin de grado

CHARACTERISTIC OPTIMALITY CRITERIA

Autor: Julio González Almeida
Tutor: Juan Manuel Rodríguez Díaz

Salamanca, 2023

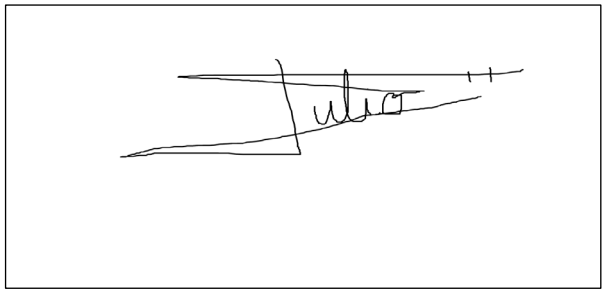
FACULTAD DE CIENCIAS

GRADO EN ESTADÍSTICA

CRITERIOS DE OPTIMIZACIÓN CARACTERÍSTICOS

Autor: Julio González Almeida

Tutor: Juan Manuel Rodríguez Díaz



RODRIGUEZ DIAZ JUAN MANUEL - 07960523Q

Firmado digitalmente por RODRIGUEZ DIAZ JUAN MANUEL - 07960523Q
Fecha: 2023.07.03 10:41:03 +02'00'

Salamanca, 2023

Certificado de los tutores TFG Grado en Estadística

D. Juan Manuel Rodríguez Díaz, profesor/a del Departamento de Estadística de la Universidad de Salamanca,

HACEN CONSTAR:

Que el trabajo titulado “Criterios de Optimización Característicos”, que se presenta, ha sido realizado por D. Julio González Almeida, con DNI 45690366-S y constituye la memoria del trabajo realizado para la superación de la asignatura Trabajo de Fin de Grado en Estadística en esta Universidad.

Salamanca, a fecha de firma electrónica.

RODRIGUEZ DIAZ JUAN MANUEL - 07960523Q

Firmado digitalmente por RODRIGUEZ DIAZ JUAN MANUEL - 07960523Q
Fecha: 2023.07.03 10:41:32 +02'00'

Fdo.: Julio González Almeida

Fdo.: Juan Manuel Rodríguez Díaz

Resumen

El trabajo se centra en la teoría del Diseño Óptimo de Experimentos, que busca seleccionar los observables de un experimento de la mejor manera posible para obtener información precisa sobre un objeto de estudio, como estimar los parámetros de un modelo de regresión. Se revisan los conceptos estadísticos y de la teoría del diseño óptimo, y se presentan los principales criterios de optimización para evaluar la calidad de un diseño.

Se destaca la importancia del Teorema de Equivalencia, que ayuda a determinar cuándo un diseño es óptimo en relación a un criterio y cuándo distintos criterios son equivalentes. También se presentan algoritmos para obtener diseños óptimos. Posteriormente, se definen los Criterios Característicos, una nueva familia de criterios de optimización basados en los coeficientes del polinomio característico de la matriz de información, que describen cada diseño. Estos nuevos criterios incluyen los conocidos D-optimización y A-optimización, que son respectivamente el determinante y la traza de la inversa de la matriz de información.

Se estudian a fondo las funciones que representan estos criterios característicos, obteniendo propiedades notables como la convexidad y diferenciabilidad. Se calcula el gradiente de los criterios, lo que permite utilizar los algoritmos para encontrar los óptimos característicos. Se aplican estos métodos a diferentes modelos conocidos y utilizados en la práctica, comparándolos con los criterios A-optimización y D-optimización. Se encuentra que los nuevos criterios intermedios son una buena alternativa para conseguir un diseño que verifique a la vez en cierta medida las propiedades de los criterios característicos más extremos.

Se destaca la generalización de los nuevos criterios a una familia más amplia que incluye otro grupo conocido, los criterios Φ_p. Además, se mencionan las interpretaciones geométricas y estadísticas de los nuevos criterios, que permiten minimizar las regiones de confianza de las estimaciones de subconjuntos de parámetros.

ÍNDICE DE CONTENIDOS

1. Introducción

1.1. Experimentos estadísticos

En torno a cualquier fenómeno que quiera ser estudiado, debido a que subyace cierta incertidumbre, el procedimiento lógico y apropiado para investigarlo es experimentar con él, de manera que puedan identificarse sus características de interés. Para poner un ejemplo, supóngase que se desea identificar el comportamiento óptimo de un sistema con respecto a su funcionamiento y su costo en distintas condiciones; entonces debe pensarse en un experimento como medio para que el sistema sea observado bajo las condiciones de interés, de tal manera que su comportamiento pueda conocerse.

A la hora de llevar a cabo un experimento, hay un aspecto muy importante pero que en muchas ocasiones, o no se tiene en cuenta o directamente tiende a infravalorarse. Estamos refiriéndonos a la formulación del problema cuya incertidumbre se desea estudiar e investigar. Si se quiere obtener una investigación que resulte exitosa, la propia investigación debe tener definido un objetivo claro y una dirección detallada con respecto al propósito del experimento. Una vez es definido el objetivo, hay que identificar la variable que se desea medir o la respuesta que se va a estudiar y, de la misma manera, todos aquellos factores que puedan influir en la posible variabilidad de las mediciones realizadas. La respuesta se conoce también como la variable dependiente; el factor o factores reciben el nombre de variables independientes y se encuentran, en cierta medida, bajo el control del investigador. Para poner un ejemplo, en una tienda o establecimiento de compras, la incertidumbre y el interés se encuentran en el número de empleados que hay disponibles, de manera que se consiga que el tiempo que están esperando los clientes no resulte excesivo. En este caso, la respuesta es el tiempo de espera y el factor que produce variabilidad el número de empleados disponible.

Como definición de lo que es el tratamiento del factor diremos que se trata de cierto valor o condición, de éste mismo, bajo el cual se llevará a cabo la medición de la respuesta correspondiente. Por ejemplo, supóngase que tenemos una gasolinera y lo que se desea es observar y, a su vez, medir el tiempo de espera cuando el establecimiento tiene a su servicio dos, cuatro o hasta seis empleados a la vez. Si el proceso de experimentación presenta varios factores que implican variabilidad en el estudio, entonces el tratamiento supondrá la combinación de los distintos niveles de cada factor; por ejemplo, si lo que queremos es estudiar el tiempo de espera a ser atendido como una función del número de empleados en el establecimiento en un determinado momento del día, entonces un tratamiento será la combinación de un número concreto de empleados en un momento concreto del día. El método que se utilizará para seleccionar el número de tratamientos será escogido a consecuencia de las metas perseguidas por el experimento. Ciertos objetivos van a requerir de unos métodos, mientras que objetivos opuestos en la investigación requerirán de métodos diferentes.

Para experimentos preliminares o, dicho de otra manera, consideraciones a tener en cuenta antes de realizar cualquier tipo de investigación científica, donde lo que se pretende es separar los principales factores de estudio, en el proceso de investigación se debe llevar a cabo la elección de los tratamientos de manera mental consiguiendo, así, una visión muy amplia

que permita de esta manera que el conocimiento extraído pueda ser utilizado por el mecanismo de estudio. En los pasos que le siguen, el experimento va a ir pasando por las distintas fases con una mayor precisión en las mediciones y cálculos con el objetivo de llegar a inferir conclusiones que sean mas concretas.

Una unidad experimental se define como el objeto (persona o cosa) que es capaz de realizar una medición de la variable respuesta después de aplicar un tratamiento dado.

Para elegir una unidad experimental o el tamaño de ésta, se dejará a elección total del experimentador. Por ejemplo, si un fabricante de baterías de coches desea comparar la duración de éstas con la de sus competidores, entonces sus baterías seleccionadas son las unidades experimentales u observaciones y el número de marcas diferentes los tratamientos. Siguiendo con los ejemplos, si se tiene interés en determinar la concentración de un residuo contaminante en un lago o en un embalse en función de la localización geográfica, entonces las localidades del lago o embalse que se seleccionan para medir la concentración del contaminante serán los tratamientos y la pequeña área superficial de cada localidad será la unidad experimental u observación correspondiente.

Cuando existe incertidumbre, los experimentos son encargados de medir y comparar las mediciones de la respuesta de las unidades experimentales esencialmente idénticas, después de que estas unidades sean expuestas a los tratamientos que han sido elegidos y aplicados por el investigador.

La influencia externa que pueda aparecer y cambiar a la medición de la variable respuesta debe ser controlada o incluso debe ser eliminada. De cualquier manera, no siempre se puede asegurar el control total de la variabilidad debida al factor externo; por ejemplo, en la práctica, casi cualquier experimento que incluye alguna actividad financiera guardará alguna interrelación con las condiciones económicas vigentes, y estas condiciones no pueden ser controladas por el investigador. Esta desviación del control de todo experimento va a estar apoyada por la reiteración de cierto experimento en una muestra de unidades experimentales para determinar la variabilidad aleatoria o el error experimental, es decir, el error que aparece debido a la aplicación del experimento. Esta se trata de la variación aleatoria que no se puede explicar, que no se puede atribuir a un cambio de tratamiento y que no puede ser controlable. Entonces, es posible inferir estadísticamente al llevar a cabo la comparativa del error experimental con las respuestas medidas producidas de la aplicación de los diferentes tratamientos.

En ciertas ocasiones, en algunas ciencias, es posible que las experimentaciones sean ideales y los resultados obtenidos sean los deseados, pero en las ciencias socioeconómicas, la variabilidad debida a las condiciones experimentales ideales tienen un lugar común ya que el ambiente no permite tener un control total y óptimo. Por ejemplo, puede llegar a ser interesante estudiar el efecto de un aumento en las tasas de interés (tratamiento) en la actividad de construcción de edificaciones inmobiliarias (respuesta) por parte de los constructores (unidades experimentales). Los tratamientos no pueden ser aplicados a las unidades experimentales y, de la misma manera, la respuesta no puede ser medida a partir de un experimento ya planeado. La información solo podrá ser registrada a medida que van cambiando las condiciones en el mundo real.

Hoy en día hay, todavía, opiniones que consideran que lo anteriormente expuesto como un ejemplo no constituye un experimento, pero de igual manera estas investigaciones científicas merecen cierta atención. Entonces, para un análisis bueno de estos datos y que ofrezca resultados con los que poder inferir conclusiones lógicas puede resultar interesante la aplicación de los distintos métodos de regresión existentes.

CANAVOS, G. C., & MEDAL, E. G. U. (1987). PROBABILIDAD Y ESTADÍSTICA (P. 651). MÉXICO: MCGRAW HILL.

1.2. Diseños estadísticos.

Los diseños estadísticos, también conocidos como diseños experimentales, son una parte fundamental de la metodología de la investigación científica. Se utilizan para planificar y organizar un estudio con el objetivo de recopilar datos de manera eficiente y obtener conclusiones estadísticamente válidas. Estos diseños se aplican en una amplia variedad de disciplinas, como la ciencia, la ingeniería, la medicina, la psicología y la agricultura, entre otras.

El diseño estadístico se refiere al método utilizado para realizar mediciones en un estudio. En general, en los experimentos diseñados estadísticamente, las unidades experimentales se seleccionan de manera aleatoria para evitar posibles sesgos sistemáticos. Este proceso aleatorio también ayuda a controlar los efectos de agentes externos no controlados por el investigador. La comparación de los tratamientos se lleva a cabo considerando que el efecto de la respuesta se enfoca únicamente en las diferencias entre los diferentes tratamientos.

En los experimentos diseñados estadísticamente, la repetición del experimento es de vital importancia, ya que nos permite medir el error asociado a la experimentación. Como mencionamos anteriormente, el tamaño del error experimental juega un papel crucial a la hora de tomar decisiones sobre si las diferencias entre los diferentes tratamientos son estadísticamente significativas o no. La repetición del experimento nos ayuda a obtener una estimación más precisa del error y a evaluar la consistencia de los resultados obtenidos. De esta manera, podemos tomar decisiones más confiables basadas en la evidencia estadística.

En el diseño de experimentos estadísticos, el objetivo principal es asignar las unidades experimentales a los tratamientos de manera adecuada para asegurar un proceso aleatorio. En este contexto, surgen dos conceptos fundamentales: el diseño completamente aleatorio y el diseño en bloques completamente aleatorio. Ambos diseños pueden utilizarse tanto en experimentos con un solo factor como en experimentos con varios factores simultáneamente.

En un diseño completamente aleatorio, la asignación de los tratamientos se realiza de manera completamente aleatoria, asumiendo que todas las unidades experimentales son homogéneas. Por lo general, la selección aleatoria se lleva a cabo mediante la generación de números aleatorios. El uso de este diseño implica que las condiciones bajo las cuales se realiza la observación de la respuesta (o cualquier otro factor bajo control del investigador) deben ser las mismas a lo largo de todo el experimento. Sin embargo, estos diseños pueden no ser adecuados cuando las observaciones se realizan en factores potenciales como el tiempo, el espacio o efectos demográficos, a menos que estos sean componentes fundamentales del experimento.

En resumen, en el diseño de experimentos estadísticos se busca asignar adecuadamente las unidades experimentales a los tratamientos, ya sea mediante un diseño completamente aleatorio o un diseño en bloques completamente aleatorio. Estos diseños garantizan un proceso aleatorio y permiten realizar inferencias estadísticas válidas en función de las condiciones y objetivos específicos del experimento.

En ocasiones, los investigadores se encuentran en situaciones donde las unidades experimentales no son todas homogéneas, lo que dificulta llevar a cabo el experimento en un entorno uniforme. En estos casos, se puede recurrir a una clasificación en la que se forman bloques homogéneos de unidades experimentales. Luego, los tratamientos se asignan aleatoriamente a las unidades dentro de cada bloque. Este enfoque se conoce como diseño en bloques completamente aleatorizado, ya que incluye todos los tratamientos y se asignan a las unidades experimentales mediante un método aleatorio.

En el proceso de investigación, es crucial agrupar las unidades experimentales en bloques, pero esto puede dar lugar a diferencias significativas en las medidas de respuesta. Estas diferencias suelen estar relacionadas con efectos espaciales, temporales o demográficos. Por ejemplo, si las unidades experimentales son seres humanos, el agrupamiento por bloques deberá considerar variables como sexo, edad, condiciones de salud y experiencia, de acuerdo con los requisitos del experimento. Si el experimento se lleva a cabo a lo largo de un período prolongado, el tiempo también se considerará como una variable en el agrupamiento por bloques.

Si los datos experimentales se recopilan en diferentes localidades o grupos, estos también se considerarán como variables para el agrupamiento por bloques. Además, si se utilizan varios instrumentos para registrar los datos, se realizará un agrupamiento de los instrumentos por bloques, incluso si son del mismo modelo, y especialmente si provienen de diferentes fabricantes.

En resumen, el agrupamiento de las unidades experimentales en bloques es esencial para la investigación, pero es importante tener en cuenta las posibles diferencias en las medidas de respuesta debido a factores espaciales, temporales o demográficos. Se deben considerar cuidadosamente las variables relevantes para el agrupamiento por bloques, ya sea sexo, edad, condiciones de salud, tiempo, ubicación o instrumentos utilizados, con el fin de controlar y comprender mejor las fuentes de variabilidad en el experimento.

En resumen, agrupar las unidades experimentales en bloques es necesario, especialmente cuando estas unidades son heterogéneas. Cuanto mayor sea la heterogeneidad, mayor será el error debido a la experimentación y menor será la probabilidad de encontrar las diferencias reales entre los tratamientos. Por lo tanto, el objetivo de los diseños estadísticos es controlar las variables que no forman parte del experimento para minimizar el error y, por ende, detectar diferencias significativas en la respuesta. Al agrupar en bloques, se busca reducir la variabilidad causada por factores externos y garantizar que las diferencias observadas en la respuesta sean atribuibles a los tratamientos y no a otros factores no controlados. Esto mejora la precisión de los resultados y facilita la identificación de las verdaderas diferencias entre los tratamientos en el estudio.

CANAVOS, G. C., & MEDAL, E. G. U. (1987). PROBABILIDAD Y ESTADÍSTICA (P. 651). MÉXICO: MCGRAW HILL.

2. Modelo lineal

El análisis de un modelo de regresión múltiple es el método por el cual se relaciona una variable dependiente (Y) y un conjunto de variables independientes ($X_1, X_2, \ldots, X_k$). A diferencia de la regresión simple, la regresión múltiple se utiliza con mayor frecuencia debido a que se ajusta mejor a situaciones reales. Esto se debe a que los fenómenos y procesos sociales suelen ser más complejos y están influenciados por un mayor número de variables, ya sea de forma directa o indirecta. La regresión múltiple nos permite considerar la interacción y el efecto conjunto de estas variables en la variable dependiente, lo que nos brinda una comprensión más completa y precisa de las relaciones subyacentes en el fenómeno que estamos estudiando. Al incluir múltiples variables independientes, podemos capturar mejor la complejidad de los fenómenos sociales y obtener una representación más realista de cómo se relacionan con la variable dependiente.

RODRÍGUEZ-JAUME, M. J., & MORA CATALÁ, R. (2001). ANÁLISIS DE REGRESIÓN MÚLTIPLE.

Sea Y una variable aleatoria que fluctúa alrededor de un valor desconocido η, esto sería

$$\mathrm{Y} = \eta + \varepsilon$$

donde ε es el error, de forma que η puede representar el valor real e Y el valor observado.

Supongamos que η toma valores distintos de acuerdo con diferentes situaciones experimentales según el modelo lineal

$$\eta = \beta_1 \mathrm{x}_1 + \ldots + \beta_\mathrm{m} \mathrm{x}_\mathrm{m}$$

donde β_i son parámetros desconocidos y x_i son valores conocidos, cada uno de los cuales representa situaciones experimentales distintas.

En general se tienen n observaciones de la variable Y. Diremos que $y_1, y_2, \ldots, y_n$ observaciones independientes de Y siguen un modelo lineal si

$$y_i = x_{i1}\beta_1 + \ldots + x_{im}\beta_m + \varepsilon_i, \qquad i = 1, \ldots, n$$

Estas observaciones de Y se pueden considerar variables aleatorias independientes y distribuidas como Y (son copias) o también realizaciones concretas (valores numéricos) para los cálculos.

La expresión del modelo lineal en forma matricial es

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{pmatrix}$$

o en forma resumida

$$\mathrm{Y} = \mathrm{X}\beta + \varepsilon$$

Los elementos que constituyen el modelo lineal son:

I. El vector de observaciones $\mathrm{Y} = (\mathrm{y}_1, \mathrm{y}_2, \acute{}\ldots, \mathrm{y}_\mathrm{n})$

II. El vector de parámetros $\beta = (\beta_1, \beta_2, \acute{}\ldots, \beta_\mathrm{m})$

III. La matriz del modelo

$$\mathrm{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix}$$

cuyos elementos son conocidos.

En problemas de regresión, la matriz X es la matriz de regresión. En los llamados diseños factoriales del Análisis de la Varianza, la matriz X recibe el nombre de matriz de diseño.

IV. El vector de errores o desviaciones aleatorias $\varepsilon = (\varepsilon_1, \varepsilon_2, \acute{}\ldots, \varepsilon_\mathrm{n})$, donde ε_i es la desviación aleatoria de y_i.

CARMONA, F. (2005). MODELOS LINEALES. *PUB. UNIV. DE BARCELONA, BARCELONA*.

2.1. Método de los mínimos cuadrados generalizados

El método de los mínimos cuadrados generalizados (MCG) es una técnica estadística utilizada para estimar los parámetros desconocidos en modelos lineales cuando los supuestos clásicos de mínimos cuadrados ordinarios (MCO) no se cumplen.

En el enfoque de mínimos cuadrados ordinarios, se asume que los errores de las observaciones son independientes, idénticamente distribuidos y siguen una distribución normal. Sin embargo, en muchas situaciones reales, estos supuestos pueden no cumplirse. Por ejemplo, los errores pueden tener una varianza no constante (heterocedasticidad) o pueden estar correlacionados (autocorrelación). En tales casos, los estimadores de MCO pueden ser ineficientes o sesgados.

El método de los mínimos cuadrados generalizados aborda estos problemas al permitir una especificación más flexible de la estructura de varianza-covarianza de los errores. En lugar de suponer que los errores son independientes e idénticamente distribuidos, el MCG permite que los errores tengan una estructura de covarianza más general.

En resumen, el método de los mínimos cuadrados generalizados es una extensión del enfoque de mínimos cuadrados ordinarios que se utiliza cuando los supuestos clásicos no se cumplen. Proporciona estimaciones robustas y eficientes de los parámetros en modelos lineales al permitir una mayor flexibilidad en la especificación de la estructura de varianza-covarianza de los errores. Esto lo convierte en una herramienta valiosa en el análisis estadístico cuando los datos presentan desviaciones de los supuestos tradicionales.

El método de los mínimos cuadrados generalizados tiene como objetivo principal estimar los parámetros de regresión β_i. Para ello, lo que hace es minimizar la siguiente función:

$$\sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2 = \sum_{i=1}^{n} \left(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i\right)^2$$

Derivando e igualando a cero quedan las siguientes ecuaciones:

$$\frac{\partial(\sum_{i=1}^{n}\hat{u}_i^2)}{\partial\hat{\beta}_1} = -2\sum_{i=1}^{n}(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) = 0$$

$$\frac{\partial(\sum_{i=1}^{n}\hat{u}_i^2)}{\partial\hat{\beta}_2} = -2\sum_{i=1}^{n}(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)X_i = 0$$

Haciendo operaciones queda lo siguiente:

$$\begin{cases} \sum_{i=1}^{n} Y_i - n\hat{\beta}_1 - \hat{\beta}_2\sum_{i=1}^{n} X_i = 0 \\ \sum_{i=1}^{n} Y_i X_i - \hat{\beta}_1\sum_{i=1}^{n} X_i - \hat{\beta}_2\sum_{i=1}^{n} X_i^2 = 0 \end{cases}$$

$$\begin{cases} \sum_{i=1}^{n} Y_i = n\hat{\beta}_1 + \hat{\beta}_2\sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} Y_i X_i = \hat{\beta}_1\sum_{i=1}^{n} X_i + \hat{\beta}_2\sum_{i=1}^{n} X_i^2 \end{cases}$$

Finalmente la estimación de los parámetros queda así:

$$\begin{cases} \hat{\beta}_2 = \dfrac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \dfrac{S_{XY}}{S_X^2} \\ \hat{\beta}_1 = \bar{Y} - \hat{\beta}_2\bar{X} \end{cases}$$

Para la estimación de los parámetros es imprescindible hacer unas ciertas suposiciones antes:

- El modelo de regresión es lineal en los parámetros

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

$$\log(Y_i) = \beta_1 + \beta_2\sqrt{X_i} + u_i$$

- Valores fijos de X o valores de X independientes del término de error: los valores que toma X pueden considerarse fijos en muestras repetidas o, en caso de ser X estocástica, haber sido muestreado junto con la variable dependiente. En ese segundo caso, $\mathrm{Cov}(\mathrm{X_i}, \mathrm{u_i}) = 0$.
- El número de observaciones debe ser mayor que el número de parámetros a estimar.
- $S_X^2 > 0$ y X no debe tener valores atípicos
- El modelo suponemos que está correctamente especificado.
- El valor medio de u_i es igual a 0: $\mathrm{E}(\mathrm{u_i}/\mathrm{X_i}) = 0$ o $\mathrm{E}(\mathrm{u_i}) = 0$ (esto último si consideramos que X es estocástica)
- Homocedasticidad o varianza constante de u_i

$$\mathrm{Var}(\mathrm{u_i}) = \sigma^2$$

- No hay autocorrelación, las observaciones se muestrean de manera independiente.

$$\mathrm{cov}(u_i, u_j / X_i, X_j) = 0, i \neq j$$

$$\mathrm{cov}(u_i, u_j) = 0, i \neq j \text{ si X es no estocástica}$$

- Los coeficientes β_i se mantendrán constantes a lo largo de toda la muestra.

Los estimadores derivados del método de los mínimos cuadrados generalizados son imparciales, consistentes y óptimos.

En primer lugar, decir que los estimadores son imparciales significa que, en promedio, no se espera que estén sesgados, es decir, que su valor estimado se acerque al valor verdadero del parámetro que se está estimando.

En segundo lugar, los estimadores son consistentes, lo que significa que a medida que aumenta el tamaño de la muestra, los estimadores convergen al valor verdadero del parámetro. Esto implica que a medida que se disponga de más datos, la precisión de los estimadores aumentará y se acercarán cada vez más al valor real.

Por último, los estimadores obtenidos a través del método de los mínimos cuadrados generalizados se consideran óptimos en el sentido de que minimizan la varianza de los errores de estimación. Esto implica que, entre todos los estimadores posibles, los estimadores de mínimos cuadrados generalizados son los más eficientes, lo que los hace deseables en términos de precisión y calidad de la estimación.

En resumen, los estimadores obtenidos mediante el método de los mínimos cuadrados generalizados son imparciales, consistentes y óptimos, lo que los convierte en estimaciones fiables y de alta calidad en el contexto de modelos lineales cuando se presentan violaciones a los supuestos clásicos de mínimos cuadrados ordinarios.

2.2. Regresión

La regresión estadística es una técnica analítica utilizada para estudiar y modelar la relación entre una variable dependiente y una o más variables independientes. Es una herramienta fundamental en el campo de la estadística y se utiliza ampliamente en diversos campos como la economía, la psicología, la biología y muchas otras disciplinas científicas.

El objetivo principal de la regresión estadística es entender cómo las variables independientes influyen en la variable dependiente y cómo se puede utilizar esta relación para predecir o estimar valores futuros de la variable dependiente. Esto se logra mediante la construcción de un modelo estadístico que represente la relación entre las variables.

En la regresión estadística, se busca encontrar los mejores estimadores de los parámetros del modelo, utilizando métodos como los mínimos cuadrados ordinarios (MCO). Estos estimadores permiten determinar la forma y la magnitud de la relación entre las variables y proporcionan información sobre la importancia relativa de cada variable independiente en la predicción o explicación de la variable dependiente.

Además de estudiar la relación lineal entre las variables, la regresión estadística también puede abordar relaciones no lineales utilizando transformaciones de variables o técnicas más avanzadas como regresión polinómica, regresión logística o regresión no paramétrica.

En resumen, la regresión estadística es una herramienta poderosa para analizar y modelar la relación entre variables. Proporciona una forma sistemática de examinar cómo los cambios en las variables independientes se relacionan con los cambios en la variable dependiente, permitiendo así la predicción y el análisis de los fenómenos y procesos en una amplia gama de disciplinas científicas.

ISGRO, C. A. (1991). ESTIMACION DE REGRESION A TROZOS (DOCTORAL DISSERTATION).

Un modelo tiene diferentes posibilidades, es decir, puede realizarse de varias formas. Estas formas serán los distintos valores de los parámetros. Cada realización del modelo se va a llamar estado denotado como θ y el conjunto de todos los estados será definido como $\ominus$.

El modelo matricial después de escoger N observaciones puede escribirse de la siguiente manera:

$$Y = X\beta + \varepsilon$$

donde

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix};\ \ \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{pmatrix},$$

El vector Y se refiere al vector formado por las observaciones y el vector ε es el de los errores producidos debido al muestreo.

Por otro lado tenemos la matriz de diseño

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{N1} & \cdots & X_{Nm} \end{pmatrix}$$

Suponiendo ademas que la esperanza de los errores es igual a cero, tenemos la matriz de covarianzas de la siguiente manera

$$\sum\nolimits_{\varepsilon} = E[\varepsilon, \varepsilon^t] = \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2 & \cdots & \sigma_{1N}^2 \\ \sigma_{12}^2 & \sigma_2^2 & \cdots & \sigma_{2N}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1N}^2 & \sigma_{2N}^2 & \cdots & \sigma_N^2 \end{pmatrix}$$

2.3. Supuestos básicos del modelo lineal

Los errores ε_i del modelo lineal se suponen como desviaciones y su comportamiento es comparado con el de las variables aleatorias verificando las condiciones de Gauss-Markov:

1) $E[\varepsilon_i] = 0 \quad i = 1, \ldots, n$
2) $var(\varepsilon_i) = \sigma^2 \quad i = 1, \ldots, n$
3) $E[\varepsilon_i \varepsilon_j] = 0 \quad \forall\, i \neq j$

Como sabemos, la condición (2) es la llamada condición de homocedasticidad del modelo y el parámetro desconocido σ^2 es la llamada varianza del modelo. La condición (3) significa que las n desviaciones son mutuamente incorrelacionadas.

Estas condiciones pueden expresarse en forma matricial como

$$\mathrm{E}[\varepsilon] = 0 \qquad \mathrm{var}(\varepsilon) = \sigma^2\, \mathrm{I_n}$$

donde $\mathrm{E}[\varepsilon]$ es el vector de esperanzas matemáticas y $\mathrm{var}(\varepsilon)$ es la matriz de covarianzas de $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$.

Si además suponemos que cada ε_i es $\mathrm{N}(0, \sigma)$ y que $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ son estocásticamente independientes, entonces diremos que el modelo definido es un modelo lineal normal. Así tendremos que

$$\mathrm{Y} \sim \mathrm{N_n}(\mathrm{X}\beta, \sigma^2 \mathrm{I_n})$$

es decir, Y sigue la distribución normal multivariante de vector de medias $\mathrm{X}\beta$ y matriz de covarianzas $\sigma^2 \mathrm{I_n}$.

Se llama rango del diseño al rango de la matriz X

$$\mathrm{r} = \mathrm{rango\ X}$$

y es un elemento muy importante en la discusión de los modelos. Evidentemente $\mathrm{r} \leq \mathrm{m}$. El valor de r es el número efectivo de parámetros del diseño, en el sentido de que si r<m es posible volver a parametrizar el modelo para que r sea igual al número de parámetros. En ciertas ocasiones el diseño comprueba que r=m, lo cual quiere decir que el rango es máximo.

Se dice que un modelo lineal se encuentra bajo las condiciones de Gauss-Markov ordinarias cuando verifica las condiciones anteriormente expuestas, a excepción de la normalidad.

CARMONA, F. (2005). MODELOS LINEALES. PUB. UNIV. DE BARCELONA, BARCELONA.

3. Diseño óptimo de experimentos

Al igual que con cualquier otro campo, la teoría del diseño óptimo de experimentos puede no ser atractiva para todos, ya que es una disciplina que demanda amplios conocimientos y desarrollos matemáticos, lo que dificulta su aplicación en la práctica. Sin embargo, se hará un esfuerzo por persuadir sobre la utilidad de esta teoría.

Aunque puede resultar desafiante, la teoría del diseño óptimo de experimentos ofrece beneficios significativos en la planificación y ejecución de estudios científicos. Al utilizar métodos de diseño óptimo, es posible obtener resultados más precisos, eficientes y confiables, maximizando así el valor de la información obtenida.

La aplicación de la teoría del diseño óptimo de experimentos permite seleccionar cuidadosamente las condiciones experimentales, distribuir adecuadamente las observaciones y controlar los factores que pueden afectar los resultados. Esto ayuda a minimizar el sesgo y la variabilidad, maximizando la capacidad de detectar y comprender las diferencias y relaciones entre las variables estudiadas.

A pesar de los desafíos asociados con su implementación, la teoría del diseño óptimo de experimentos se considera una herramienta valiosa para los investigadores y científicos que buscan obtener resultados de alta calidad y tomar decisiones basadas en evidencias sólidas. A través de un enfoque riguroso y sistemático, esta disciplina puede mejorar la validez y la eficiencia de los estudios experimentales, impulsando así el avance del conocimiento en diversas áreas.

En resumen, aunque la teoría del diseño óptimo de experimentos puede ser compleja y requerir un conocimiento profundo, su aplicación puede aportar beneficios sustanciales en términos de precisión, eficiencia y confiabilidad de los resultados experimentales. Es una disciplina que busca mejorar la calidad de la investigación científica y promover la toma de decisiones fundamentadas en datos sólidos.

3.1. Crítica al diseño óptimo de experimentos

Primeramente hay que ver diversos aspectos negativos o inconvenientes de la teoría del diseño óptimo de experimentos:

- Elección del modelo final sin tener anteriormente las observaciones. Esto hace que el resultado obtenido, es decir, el diseño óptimo obtenido dependa en un alto grado del modelo elegido.
- Dentro de la teoría para los modelos no lineales, cabe destacar que hay que tener en cuenta que se deben escoger ciertos valores iniciales para los diferentes parámetros del modelo.
- Decidir si un diseño es el óptimo o no siempre dependerá de un criterio que, a su vez, va a depender del estudio estadístico. Y, aquí, de nuevo se va a elegir el modelo antes de escoger las observaciones. El problema que surge en este aspecto, es que un diseño

puede resultar muy bueno con cierto criterio, pero a la vez puede estar dejando de lado algún aspecto que resultaría de gran importancia en el estudio.

- Los diseños exactos, son mas complejos matemáticamente pero en la práctica se pueden llevar a cabo, mientras que los diseños aproximados son mas fáciles de obtener debido a la existencia del teorema de equivalencia. Pero estos diseños en la práctica podrían resultar poco eficientes en el caso que los tamaños de muestra sean bajos.
- En ciertas ocasiones, estos diseños requieren de la obtención de observaciones "extremas". El problema que surge aquí es que obtener estas observaciones no se podría realizar por razones tanto prácticas, como en muchas ocasiones por razones éticas.
- Puede ocurrir que el diseño obtenido sea muy bueno para algún parámetro pero malo para el resto.

A partir del estudio de esta serie de inconvenientes, es fácil llegar a la conclusión de que el problema que surge no radica en la teoría del diseño óptimo de experimentos sino que aparece en los modelos matemáticos.

- Para la elección del modelo, aunque no contemos de antemano con los datos, siempre va a haber cierta experiencia del investigador, estudios anteriores en los que fijarse... Además, la elección del modelo se realizará entre un grupo reducido de ellos.
- Para el caso del uso de modelos no lineales y la elección de ciertos valores iniciales, existen estudios de robustez y sensibilidad.
- En cuanto a la elección del criterio para la obtención del diseño óptimo, existe la posibilidad de elegir criterios combinados que den buenos resultados para diversos puntos de vista.
- El cálculo de diseños exactos se dedicará solamente para tamaños de muestra bajos mientras que para tamaños de muestra altos se considerarán los diseños aproximados, donde existen algoritmos implementados que permiten tratar con gran cantidad de datos y a su vez ofreciendo resultados buenos y eficientes. Además, están utilizando métodos numéricos cuyas soluciones se aproximan a las soluciones reales.
- En cuanto a la elección de observaciones "extremas", existe la posibilidad de calcular diseños restringidos o de definir de otra manera los límites del dominio experimental de modo que las condiciones frontera no sean tan extremas y que a su vez sean asequibles en la práctica.
- En cuanto al problema de la invarianza por cambios de escala, existen diseños igualmente eficientes para todos los parámetros del modelo.

3.2. Conceptos básicos.

El primer paso en cualquier investigación es definir y especificar el modelo que se utilizará. Una vez que se ha establecido el modelo, el siguiente paso es determinar las condiciones experimentales en las cuales se medirá la respuesta, con el objetivo de obtener una inferencia estadística clara y al menor costo posible.

Para lograr esto, se lleva a cabo un diseño experimental en el cual se eligen cuidadosamente los niveles de las variables independientes y se determina el tiempo entre las mediciones de la respuesta. Estas decisiones se basan en criterios de optimalidad que tienen un significado estadístico.

El objetivo del diseño experimental es maximizar la eficiencia de la recopilación de datos, asegurando que la información obtenida sea estadísticamente válida y relevante. Esto implica encontrar el equilibrio adecuado entre la precisión de las mediciones y los recursos disponibles, como el tiempo y los costos asociados.

En resumen, el diseño experimental busca garantizar una inferencia estadística clara y eficiente mediante la elección cuidadosa de las condiciones experimentales. Se basa en criterios de optimalidad que permiten obtener conclusiones válidas con el mínimo costo posible. Este enfoque garantiza que la investigación sea rigurosa y que los resultados sean confiables y significativos desde el punto de vista estadístico.

Unido al modelo lineal explicado anteriormente, se define el siguiente diseño aproximado,

$$\xi = \begin{pmatrix} \mathrm{x}_1 & \cdots & \mathrm{x}_\mathrm{n} \\ \omega_1 & \cdots & \omega_\mathrm{n} \end{pmatrix}$$

con $\omega_i = \xi(x_i)$ como una medida de probabilidad definida en B, conjunto de Borel de X que incluye los conjuntos unitarios; tal que ξ tiene soporte finito. El soporte de ξ es $\mathrm{Supp}(\xi) = [\mathrm{x}_1, \ldots, \mathrm{x}_\mathrm{n}]$, n: número de puntos de soporte de ξ, y las observaciones $\mathrm{Y}(\mathrm{x})$ se hacen en $x_1, \ldots, x_n$ con frecuencias (o pesos) aproximadamente proporcionales a $\omega_1, \ldots, \omega_n$.

Para la formación de un diseño que sea óptimo lo primero es elegir la matriz que contenga los puntos candidatos a pertenecer al diseño óptimo y que se conoce como matriz de puntos candidatos y está formada por N filas siendo denotada de la siguiente manera:

$$\xi_N = \begin{pmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ \vdots & \vdots \\ 1 & 1 \end{pmatrix}$$

Por otro lado, y a partir de la matriz de puntos candidatos, está la matriz de diseño X que contiene n filas y con p coeficientes y se escribirá de la siguiente manera:

$$\mathrm{X} = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 1 \end{bmatrix}$$

La primera columna de la matriz de diseño representa el término constante de β_0 que es la intersección, y el resto de las columnas representan los términos de un modelo de regresión lineal, siendo las variables x_i las que actúan como variables regresoras permitiendo hacer estimaciones sobre la variable respuesta.

De esta manera, para hacer la selección de los mejores puntos candidatos el proceso se realiza mediante un adecuado criterio de selección. La selección del mejor grupo de puntos

candidatos se conoce como optimalidad y la matriz de diseño correspondiente es llamada matriz de diseño óptimo.

NARANJO-PALACIOS, F., RIOS-LIRA, A. J., PANTOJA-PACHECO, Y. V., & TAPIA-ESQUIVIAS, M. (2020). DISEÑOS ORTOGONALES DE TAGUCHI FRACCIONADOS. INGENIERÍA, INVESTIGACIÓN Y TECNOLOGÍA, 21(2).

Para cada diseño ξ se define la matriz de información de Fisher de acuerdo con las siguientes condiciones:

- Si en el modelo de regresión los errores tienen la siguiente estructura $\varepsilon \sim N_N(0, \sigma^2 I)$ la matriz de información de Fisher está dada por la siguiente igualdad:

$$M(\xi, \theta) = \sum_{i=1}^{k} f(t_i, \theta) f^T(t_i, \theta) w_i$$

donde $f(t_i, \theta) = \left[\frac{\partial \eta(t_i,\theta)}{\partial \theta_1}, \ldots, \frac{\partial \eta(t_i,\theta)}{\partial \theta_p}\right]^T$; para $i = 1,2,\ldots,k$

en este caso la matriz de información depende del vector de parámetros θ.

- Si en el modelo de regresión no lineal los errores tienen la siguiente estructura $\varepsilon \sim N_N(0, \sum_\beta)$ y $\xi = \{t_1, t_2, \ldots, t_N\}$ es un diseño exacto de N-puntos, entonces la matriz de información de Fisher está dada por:

$$M(\xi, \theta, \beta) = \begin{pmatrix} M_{11} & 0 \\ 0 & M_{22} \end{pmatrix}$$

donde las matrices M_{11} y M_{22} se definirán a continuación

$$M_{11} = F^T(\theta) \sum_{\beta}^{-1} F(\theta)$$

$$M_{22} = \left[\frac{1}{2} tr\left(\sum\nolimits_{\beta}^{-1} \frac{\partial \sum_\beta}{\partial \beta_k} \sum\nolimits_{\beta}^{-1} \frac{\partial \sum_\beta}{\partial \beta_I}\right)\right]_{qxq}$$

$$F(\theta) = \left[\frac{\partial \eta(t_i, \theta)}{\partial \theta_j}\right]_{Nxp}$$

y

$$\sum\nolimits_{\beta} = \left[Cov(\varepsilon_i, \varepsilon_j)\right]_{NxN} = \left[E\left(\varepsilon(t_i)\varepsilon(t_j)\right)\right]_{NxN}$$

La matriz de información de Fisher es una matriz particionada, en donde la diagonal principal de esta matriz está conformada por submatrices cuadradas que dependen de los vectores θ, β y ξ.

Para cada diseño ξ se define la matriz de momentos:

$$M(\xi) \equiv \int_x f(x) f^T(x) d\xi(x) = \sum_{i=1}^{n} f(x_i) f^T(x_i) \omega_i$$

Saber la cantidad de información que nos aporta la matriz de momentos va a depender de los criterios de optimalidad, que son aquellos capaces de maximizar algún funcional real (con significado estadístico) de la matriz de momentos sobre Ξ; clase de todos los diseños definidos en $\mathcal{B}$. Estos criterios de optimalidad se presentan a continuación, siguiendo el enfoque de Pukelsheim (1993), quien introduce la matriz de información C_k, función puente que da cuenta de la "información" contenida en combinaciones lineales de θ; $C_k$es una función del conjunto de las matrices definidas no negativas de orden m, $\mathrm{N\,N\,D\,(m)}$, en el conjunto de las matrices simétricas de orden q, $\mathrm{Sim(q)}$. Concluyendo con la noción de función de información ∅. La matriz de información intuitivamente mide la información que aporta el sistema de parámetros K^T ⊖, mientras que la función de información la cuantifica por medio de un número real.

Se considera el caso general, cuando el investigador está interesado en la estimación de q-combinaciones lineales de θ. Es decir, la estimación del subsistema $K^T\theta$, donde $\mathrm{K} \in \mathbb{R}^{\mathrm{mxq}}$ es conocida y $\mathrm{r(K)} = \mathrm{q}$.

- Sea ξ un diseño factible para $K^T\theta$, es decir, $\mathrm{C(K)} \subseteq \mathrm{C(M(\xi))}$, $\mathrm{C(A)}$ es el espacio generado por las columnas de la matriz A. Se define la matriz de información como la función:

$$C_K: NND(m) \longrightarrow Sim(q)$$

tal que: $C_K\big(M(\xi)\big) = (K^T M^- K)^{-1}$, A^-denota una inversa generalizada de A.

Por notación, $\mathrm{A} \geq 0$ si y solo si $\mathrm{A} \in \mathrm{NND(m)}$; $\mathrm{A} \geq \mathrm{B}$ si y solo si $\mathrm{A} - \mathrm{B} \geq 0$.

- Si $\mathrm{K} = \mathrm{I_m}$, interesa estimar θ, y $M(\xi)$ es no singular, entonces $C_I\big(M(\xi)\big) = M(\xi)$. Es decir, la matriz de información coincide con la matriz de momentos; por esta razón en la literatura M también se llama matriz de información.

La cuantificación de la información suministrada para cada diseño, ya sea por la matriz de momentos o la matriz de información, es definida a partir de una función de valor real Φ.

Para lo que sigue Φ, denotará una función de información.

Existen distintos tipos de diseños sobre los que se desarrolla toda una teoría adaptada a su tipología:

Se denomina diseño discreto o diseño exacto de N observaciones a una sucesión de N puntos de χ, $\{x_1, \ldots, x_N\}$, donde podría haber repeticiones.

Un diseño exacto se refiere a un diseño experimental en el cual se busca evaluar y analizar todas las posibles combinaciones de niveles de los factores en estudio. Esto implica que se consideran todas las posibles variaciones de los factores y se asigna cada combinación a una unidad experimental. En otras palabras, se busca cubrir todas las posibles configuraciones de los factores en el diseño.

Considerando diseños exactos, el problema de diseño consistirá en elegir N observaciones que minimicen un funcional apropiado de $\mathrm{Cov}(\widehat{\theta})$. Por cuestión de conveniencia, se considerará la matriz de información, que definida para diseños exactos es:

$$M(\xi_N) = \frac{1}{N} X_N^t X_N = \frac{1}{N} \sum_{i=1}^{N} f(x_i) f^t(x_i)$$

Dadas N observaciones, debemos decidir cuantas N_i debemos tomar en x_i, con $\sum_{i=1}^{k} N_i = N$ de modo que la matriz $X^t X$ sea tan grande como sea posible en algún sentido. Obsérvese que si $X_N^t X_N$ es no degenerada, el problema de minimizar cierta función de $\mathrm{Cov}(\widehat{\theta})$ es equivalente, salvo constante, a maximizar la matriz de información.

Por otro lado, un diseño aproximado, también conocido como diseño fraccional o submuestreo, se utiliza cuando no es posible o práctico evaluar todas las combinaciones posibles de niveles de los factores. En este caso, se selecciona una fracción de las combinaciones posibles para ser evaluadas en el diseño experimental. Esto implica que solo se estudian un subconjunto de todas las combinaciones posibles, lo que permite reducir el número de unidades experimentales necesarias y, por lo tanto, el costo y el tiempo requeridos para el experimento.

Cualquier medida de probabilidad, ξ, sobre χ con soporte en un conjunto finito, es un diseño aproximado o asintótico.

Es decir, si el diseño ξ, pone peso $\xi(x_i)$ en $x_i, i = 1, \ldots, k$. y el número total de observaciones se realizan es N, entonces, se tomarán aproximadamente N. $\xi(x_i)$ observaciones en x_i, $i = 1, \ldots, k$. Eligiendo un N suficientemente grande, es posible aproximar un diseño de estas características a uno exacto.

El diseño concentrado en los puntos $x_1, \ldots, x_k$ con pesos respectivos $\omega_1, \ldots, \omega_k$ $\left(0 \leq \omega_i \leq 1 \; para \; i = 1, \ldots, k; \; \sum_{i=1}^{k} \omega_i = 1\right)$ se denotará por

$$\xi = \begin{Bmatrix} x_1 & \cdots & x_k \\ \omega_1 & \cdots & \omega_k \end{Bmatrix}$$

siendo $\xi(x_i) = \omega_i$. Por tanto, la matriz de información para diseños aproximados se define como

$$M(\xi) = \sum_{x \in \chi} \sigma^{-2} f(x) f^t(x) \xi(x)$$

GARCÍA-CAMACHA GUTIÉRREZ, I. (2018). DISEÑO ÓPTIMO DE EXPERIMENTOS PARA MODELOS DE MEZCLAS APLICADOS EN LA INGENIERÍA Y EN LAS CIENCIAS EXPERIMENTALES.

Cabría aún una definición mas general del diseño como una medida de probabilidad cualquiera, en cuyo caso aparecerían también diseños continuos. Estos diseños, en ciertas ocasiones, no son viables en la práctica pero pueden llegar a ser convenientes para demostrar algunas propiedades. Además, eligiendo n suficientemente grande, podremos aproximar un diseño de estas características a uno exacto, tomando un número cercano a $n\xi(x)$ observaciones en el punto x.

RODRÍGUEZ HERNÁNDEZ, M. D. L. M. (2016). DISEÑO ÓPTIMO DE EXPERIMENTOS PARA MODELOS DE ADAIR Y MEZCLA DE DISTRIBUCIONES.

En el desarrollo de la teoría del Diseño Óptimo de experimentos es necesario "enriquecer" el conjunto de posibles diseños experimentales a considerar con el fin de que sea un conjunto convexo.

Se define un diseño general continuo con n puntos diferentes como las n parejas:

$$\begin{Bmatrix} x_1 & \dots & x_n \\ w_1 & \dots & w_n \end{Bmatrix} \; donde \; \sum w_j = 1$$

Siendo $w_j > 0$ expresando el "peso relativo" en el diseño.

Se puede ver claramente que los diseños exactos son un caso particular de los diseños continuos si los expresamos como:

$$\begin{Bmatrix} x_1 & \dots & x_n \\ r_1 & & r_n \\ \overline{N} & \dots & \overline{N} \end{Bmatrix}$$

y haciendo $w_j = \frac{r_j}{N}$ el peso relativo del punto x_j en el diseño. El inverso no es necesariamente cierto, puesto que los pesos w_j pueden no ser racionales.

Se utiliza la letra griega δ para simbolizar un diseño general.

En definitiva, la diferencia fundamental entre un diseño exacto y un diseño aproximado radica en la cobertura de todas las posibles combinaciones de niveles de los factores. Mientras que un diseño exacto busca evaluar todas las combinaciones, un diseño aproximado selecciona solo una fracción de las combinaciones para su estudio. Esto puede tener implicaciones en la precisión y la generalización de los resultados obtenidos. Sin embargo, los diseños aproximados son una herramienta útil cuando se busca reducir los recursos necesarios sin comprometer completamente la validez del estudio.

PAGURA, J. A., & PUIGSUBIRÁ, C. (2008). DISEÑO ROBUSTO DE PARAMETROS. ALTERNATIVAS BASADAS EN DISEÑO ÓPTIMO.

3.3. Criterios de optimización

3.3.1. Introducción

En el diseño de experimentos, los criterios de optimización son principios o reglas que se utilizan para seleccionar la configuración óptima de las condiciones experimentales. Estos criterios buscan maximizar la eficiencia de la recopilación de datos y garantizar la calidad y la validez de los resultados obtenidos.

Un buen diseño experimental debería conducir a estimaciones precisas de los parámetros y a varianzas de predicción reducidas en todos los puntos de la región experimental E_x.

Existen varios criterios de optimización utilizados en el diseño de experimentos, y la elección del criterio adecuado depende de los objetivos y las restricciones específicas del estudio. Algunos de los criterios de optimización más comunes incluyen:

I. Criterio D: Este criterio se basa en minimizar la varianza de los estimadores de los parámetros del modelo. El objetivo es obtener estimaciones precisas y confiables de los efectos de los factores en estudio.

II. Criterio A: Este criterio se centra en maximizar la capacidad de discriminación del diseño. Busca maximizar la detección de diferencias significativas entre los tratamientos o condiciones experimentales.

III. Criterio G: Este criterio se enfoca en minimizar la correlación entre los estimadores de los parámetros del modelo. El objetivo es reducir la multicolinealidad y mejorar la precisión de las estimaciones.

IV. Criterio I: Este criterio busca maximizar la información mutua entre los parámetros del modelo y los datos observados. Se utiliza en el diseño de experimentos con restricciones de información y tiene como objetivo maximizar la ganancia de información.

Estos son solo algunos ejemplos de criterios de optimización utilizados en el diseño de experimentos. Cada criterio tiene sus propias ventajas y limitaciones, y su elección dependerá de los objetivos y las características específicas del estudio. En general, los criterios de optimización ayudan a diseñar experimentos eficientes y efectivos, maximizando la información obtenida y garantizando la validez de los resultados estadísticos.

PAGURA, J. A., & PUIGSUBIRÁ, C. (2008). DISEÑO ROBUSTO DE PARAMETROS. ALTERNATIVAS BASADAS EN DISEÑO ÓPTIMO.

A la hora del diseño y realización de un experimento o una serie de experimentos con el objetivo de obtener un modelo, se va a distinguir lo que se denomina como región experimental, que se refiere al conjunto de valores posibles de las variables $x_1, \ldots, x_n$ y las combinaciones posibles de estos. Por otro lado nos encontramos con la matriz de diseño o matriz X y la matriz de información que es referida a $\acute{X}X$ y tiene dimensión pxp. Se dice que un diseño es exacto cuando se le da una probabilidad de $\frac{1}{n}$ a cada uno de los n puntos (estos puntos no tienen por qué ser diferentes de la región experimental). En forma teórica y para la literatura, los diseños exactos pueden ser considerados como una versión discreta de la literatura general. En ellos aparece la medida conocida como ξ definida sobre la región experimental x centrando la atención del estudio en la matriz de información $\mathrm{M}(\xi)$.

En la teoría del diseño de experimentos, existe una amplia literatura que aborda los diseños aproximados, con un enfoque riguroso desde el punto de vista matemático. Sin embargo, en la práctica, no siempre es posible obtener un diseño exacto debido a que la proporción de observaciones necesarias en un punto específico de la región experimental puede ser un número irracional.

En este contexto, es relevante mencionar el concepto de "cápsula" de las variables independientes, que se refiere al conjunto convexo más pequeño que contiene todos los puntos de diseño. También se puede definir como la intersección de los conjuntos convexos que incluyen a todos los puntos del diseño. En este trabajo, trabajaremos con matrices de diseño, que son equivalentes a los puntos seleccionados de la región experimental donde se llevarán a cabo los experimentos.

Los criterios de optimalidad se dividen en dos clases principales: aquellos que se centran en la varianza de los estimadores de los parámetros (criterios D, A y E) y aquellos que se centran en la varianza de los estimadores de la variable respuesta (criterios V y G). En este trabajo, nos enfocaremos específicamente en los criterios de D-optimalidad, A-optimalidad y E-

optimalidad, que buscan minimizar la varianza de los estimadores de los parámetros en el diseño del experimento.

En resumen, aunque existe una teoría rigurosa sobre diseños aproximados, en la práctica no siempre es posible obtener un diseño exacto. En este trabajo, nos centraremos en los criterios de optimalidad D, A y E, que buscan minimizar la varianza de los estimadores de los parámetros en el diseño del experimento.

SEIER, E. (1990). MATRIZ DE PROYECCIÓN Y DISEÑOS OPTIMOS PARA REGRESIÓN. PRO MATHEMATICA, 4(7-8), 31-69.

Como se ha mencionado previamente, el objetivo principal del diseño de experimentos es encontrar un diseño que minimice la varianza tanto como sea posible, lo cual conduce a una maximización de la información obtenida. Sin embargo, no existe un criterio único que garantice el mejor diseño en todos los casos. Los criterios para elegir el diseño óptimo dependerán de la situación y los objetivos específicos del experimento, así como de la facilidad de cálculo y otros aspectos más subjetivos.

En este contexto, es importante definir el concepto de función criterio, la cual se utiliza para evaluar y comparar diferentes diseños de experimentos. La función criterio tiene como objetivo cuantificar la calidad de un diseño en función de sus propiedades estadísticas. Algunas de las propiedades más relevantes de la función criterio son la capacidad de proporcionar medidas de eficiencia, la capacidad de evaluar diseños en términos de precisión y la capacidad de considerar restricciones adicionales del experimento.

En resumen, la elección del diseño óptimo en el diseño de experimentos depende de múltiples factores y criterios específicos que se adaptan a los objetivos y condiciones del experimento en cuestión. La función criterio desempeña un papel fundamental al proporcionar una medida cuantitativa de la calidad del diseño y considerar propiedades estadísticas relevantes.

DÍAZ, J. M. R. (2001). CRITERIOS CARACTERÍSTICOS EN DISEÑO ÓPTIMO DE EXPERIMENTOS (VOL. 67). UNIVERSIDAD DE SALAMANCA.

3.3.2. Funciones criterio

Las funciones criterio son herramientas utilizadas en el diseño de experimentos para evaluar y comparar diferentes diseños en función de su eficiencia y capacidad para proporcionar información precisa y confiable. Estas funciones se utilizan como guías para la selección del diseño óptimo, es decir, aquel que cumple con los objetivos del experimento de la manera más efectiva y eficiente posible.

El problema de diseño para el sistema parametral $K^T\theta$ consiste en encontrar un diseño ξ^*que sea factible y que minimice la función criterio.

Por las propiedades de ϕ y C_K, principalmente la semicontinuidad superior y la compacidad de χ, el máximo anterior se alcanza para algún diseño ξ.

A continuación, se presenta una clase de funciones de información, denominada matriz de medias (matrix means), que incluyen los criterios de optimalidad de mayor popularidad.

Sea $C \epsilon NND(q), \text{para } C > 0$:

$$\phi_p(C) = \begin{cases} \lambda_{max}(C), & p=\infty \\ \left[\frac{1}{q}tr(C^p)\right]^{\frac{1}{p}}, & p\neq 0, p\neq \pm\infty \\ (det(C))^{\frac{1}{q}}, & p=0 \\ \lambda_{min}(C), & p=-\infty \end{cases}$$

A continuación, se realizarán una serie de anotaciones a cuestión de los criterios ϕ_p-óptimos.

- ϕ_p es una función de información para $p\epsilon[-\infty, 1]$.
- Si un diseño ξ maximiza el criterio anterior, se dice que el diseño es ϕ_p -óptimo $(p\epsilon[-\infty, 1])$.
- Si $C = C_K\big(M(\xi)\big) = (K^T M(\xi)^- K)^{-1}$ y $p\epsilon\{0, -1, \infty\}$ se tienen los criterios de optimalidad mas populares que dependen de la maximización del respectivo funcional evaluado en la matriz de información (o en algunos casos evaluados en la matriz de diseño); ellos son, respectivamente,

CRITERIO D-ÓPTIMO

El criterio D-optimalidad es un método utilizado para hallar diseños que permite estimar de manera óptima los parámetros del modelo, este criterio se basa en minimizar el volumen del elipsoide de confianza, asociado a la estimación de máxima verosimilitud de los parámetros del modelo. Un diseño D-óptimo es aquel que proporciona la máxima precisión en la estimación de los parámetros.

Este criterio se apoya en el cálculo del determinante de la matriz de información. El determinante de una matriz es un único número, por lo tanto resulta de realizar una serie de operaciones con los elementos de la matriz de información. Para el cálculo del determinante de una matriz es imprescindible que dicha matriz sea cuadrada. Se obtiene de restar la multiplicación de los elementos de la diagonal principal de la matriz y la multiplicación de los elementos de la diagonal secundaria de la misma matriz.

El determinante tiene grandes utilidades que son usadas en gran cantidad de campos y uno de ellos es que nos sirve de criterio para la obtención de diseños óptimos en modelos estadístico

El criterio D-optimalidad se define como el siguiente funcional asociado a un diseño:

$$\Phi(\xi, \theta) := \Phi_D\big(M(\xi; \theta)\big) = |M^{-1}(\xi; \theta)|^{\frac{1}{p}}$$

donde p: es el número de parámetros del modelo.

El diseño que resulta de maximizar la anterior expresión se conoce como diseño D-óptimo.

El criterio D-optimalidad, cuando los errores estan correlacionados, se define como el siguiente funcional:

$$\Phi\big(M(\xi; \theta, \beta)\big) = |M(\xi; \theta, \beta)|^{\frac{1}{(p+q)}} = \{|M_{11}(\xi; \theta)||M_{22}(\xi; \beta)|\}^{\frac{1}{(p+q)}}$$

Un diseño ξ^* es D-óptimo si maximiza la expresión anterior, es decir, equivale a tener un diseño que minimice lo siguiente:

$$-\log|\mathrm{M}(\xi;\theta,\beta)| = -\log|\mathrm{M}_{11}(\xi;\theta)| - \log|\mathrm{M}_{22}(\xi;\beta)|$$

Con Φ un funcional real y $|M|$ denotando el determinante de la matriz M es mas útil el uso de funciones convexas como el logaritmo natural.

MOSQUERA-BENÍTEZ, J. C., & LÓPEZ-RÍOS, V. I. (2019). EFECTO DE DISTRIBUCIONES A PRIORI EN LOS DISEÑOS D-ÓPTIMOS BAYESIANOS PARA UN MODELO NO LINEAL CORRELACIONADO. CIENCIA EN DESARROLLO, 10(2), 165-170.

El criterio D-óptimo verifica que:

I. Φ_D es continua en M

II. La función Φ_D es convexa en M y estrictamente convexa en M_+.

III. En las matrices en que Φ_D es finita, también es diferenciable. Ademas su gradiente es:

$$\nabla[-\log \det \mathrm{M}] = -\mathrm{M}^{-1}$$

El criterio D-óptimo es especialmente útil cuando el interés principal radica en la estimación precisa de los parámetros del modelo. Sin embargo, es importante tener en cuenta que este criterio no tiene en cuenta directamente la interpretabilidad o la facilidad de implementación del diseño, ya que se centra exclusivamente en la eficiencia estadística.

El criterio D-óptimo tiene una ventaja significativa en comparación con otros criterios debido a su facilidad de cálculo. Además, se puede interpretar fácilmente desde un punto de vista geométrico. Esto se debe a que las longitudes de los ejes del elipsoide de confianza de las estimaciones de los parámetros del modelo están relacionadas con las raíces cuadradas de los valores propios de la matriz de covarianzas. En otras palabras, al minimizar el determinante de la matriz de información, un diseño D-óptimo también minimiza el volumen de la región de confianza de los parámetros.

Esta interpretación geométrica permite comprender intuitivamente cómo un diseño D-óptimo afecta la precisión de las estimaciones de los parámetros. Al reducir el volumen de la región de confianza, el diseño garantiza una mayor precisión en las estimaciones, lo que significa que los valores estimados de los parámetros serán más cercanos a los valores verdaderos

También aparece este criterio de la siguiente manera:

$$\Phi[\mathrm{M}(\xi)] = \det\left[\mathrm{M}^{-\frac{1}{\mathrm{m}}}(\xi)\right]$$

Así, es fácil conseguir la homogeneidad necesaria, que permite un buen uso de la eficiencia. El logaritmo consigue que dar una prueba mas fácil de la convexidad y a la vez obtener un gradiente mas simple.

RODRÍGUEZ HERNÁNDEZ, M. D. L. M. (2016). DISEÑO ÓPTIMO DE EXPERIMENTOS PARA MODELOS DE ADAIR Y MEZCLA DE DISTRIBUCIONES.

CRITERIO A-ÓPTIMO

Este criterio busca minimizar la varianza promedio de los estimadores de los parámetros. No toma en cuenta las covarianzas de los estimadores. Utilizando la $\frac{1}{p}\sum_{j=1}^{p} Var(\widehat{\beta_j}) = \frac{\sigma^2}{p}\sum_{j=1}^{p} c_{jj} = \frac{\sigma^2}{p} tr(\acute{X}X)^{-1}$

Asumiendo la varianza σ fija, el criterio de A-optimalidad busca minimizar la traza de $(\acute{X}X)^{-1}$. En John (1972) se demuestra que para minimizar esta traza los experimentos deben realizarse en puntos a la misma distancia del centro.

Para entender lo arriba expuesto hay que tener el cuenta lo que es la traza de una matriz, en este caso de la matriz de información. A continuación se detalla brevemente con una definición:

La traza de una matriz de información, que imprescindiblemente debe de ser cuadrada, es un único número al igual que el determinante pero solo se tendrán en cuenta los elementos de la diagonal principal. La traza se obtiene realizando la suma de todos los elementos de la diagonal.

SEIER, E. (1990). MATRIZ DE PROYECCIÓN Y DISEÑOS OPTIMOS PARA REGRESIÓN. PRO MATHEMATICA, 4(7-8), 31-69.

Este criterio equivale a minimizar la suma o promedio de las varianzas de los estimadores de los parámetros. El factor $-\frac{1}{\mathrm{m}}$ se considera para que la función sea homogénea positiva. Entonces se comprueba que:

$$\sum_{j=1}^{m} Var(\widehat{\theta_j}) \propto Tr(M^{-1}(\xi))$$

El criterio A-óptimo es especialmente útil cuando se tiene interés en la precisión global de las estimaciones de los parámetros y no se da una importancia significativa a estimaciones individuales. Al minimizar la traza de la matriz de covarianzas inversa, se logra un diseño que equilibra la precisión de todas las estimaciones de los parámetros.

En resumen, el criterio A-óptimo se basa en minimizar la traza de la matriz de covarianzas inversa para obtener un diseño que mejore la precisión global de las estimaciones de los parámetros. Es un criterio útil cuando se busca equilibrar la precisión de todas las estimaciones en lugar de enfocarse en una estimación específica.

Los diseños A-óptimos se consideran invariantes con respecto a reparametrizaciones ortogonales.

ARROYO, C. (2022). DISEÑO ÓPTIMO DE EXPERIMENTOS: APLICACIONES PARA LA ECUACIÓN DE ANTOINE, DISEÑOS D-AUMENTADOS Y PAQUETE OPTEDR.

C OPTIMALIDAD

- c-optimalidad. Si $\mathrm{K} = \mathrm{c}, \mathrm{c}\epsilon\mathbb{R}^{\mathrm{mx1}}$ entonces el criterio asociado se denomina c-optimalidad; se puede mostrar que la única función de información es la identidad: $\phi(\delta) = \delta$ y el problema de diseño se reduce a encontrar un diseño ξ^* que sea factible para $c^T\theta$ y maximice la función de información:

$$\phi\left(C_c\left(M(\xi)\right)\right) = C_c\left(M(\xi)\right) = (c^T M(\xi)^- c)^{-1}$$

observese que el lado derecho representa el inverso de la varianza asociada al estimador óptimo para $c^T\theta$; luego los diseños c-óptimos son aquellos que minimizan la varianza de $c^T\hat{\theta}$.

LÓPEZ, V. I., & RAMOS, R. (2007). UNA INTRODUCCIÓN A LOS DISEÑOS ÓPTIMOS. REVISTA COLOMBIANA DE ESTADÍSTICA, 30(1), 37-51.

El diseño óptimo requiere de una función que pueda emplearse como una medición de la calidad de la inferencia resultante. A esta función se le va a denominar como función criterio.

NARANJO-PALACIOS, F., RIOS-LIRA, A. J., PANTOJA-PACHECO, Y. V., & TAPIA-ESQUIVIAS, M. (2020). DISEÑOS ORTOGONALES DE TAGUCHI FRACCIONADOS. INGENIERÍA, INVESTIGACIÓN Y TECNOLOGÍA, 21(2).

La función criterio se va a definir de la siguiente manera:
Diremos que una función

$$\Phi\colon\ \mathcal{M} \longrightarrow \mathbb{R} \cup \{+\infty\}$$

acotada inferiormente es una función criterio si cumple lo siguiente:

$$\mathrm{Var}_\xi g \leq \mathrm{Var}_\eta g \ \text{ para todo funcional } \ g \Rightarrow \Phi[M(\xi)] \leq \Phi[M(\eta)]$$

Entonces se trata de un criterio de Φ-optimización. Un diseño capaz de minimizar $\Phi[M(\xi)]$ se denominará como diseño Φ-óptimo.

Una vez definido el concepto de función criterio se van a detallar las propiedades de la misma.

Toda función criterio Φ, tal y como ha sido definida con anterioridad es siempre decreciente, de la siguiente manera:

$$M(\xi) \geq M(\eta) \Rightarrow \Phi[M(\xi)] \leq \Phi[M(\eta)]$$

Entonces se cumple lo siguiente

$$Var_\xi g \leq Var_\eta g \ para\ todo\ funcional\ g$$

Es importante destacar que existen casos en los que varias funciones criterio pueden llevar a la selección del mismo diseño óptimo. Esto significa que diferentes funciones criterio pueden producir diseños óptimos idénticos. Este concepto se demuestra mediante el teorema de equivalencia.

Es evidente que si un diseño es óptimo, cualquier diseño asociado con la misma matriz de información también será óptimo. Sin embargo, también es posible que diferentes matrices de información conduzcan a diseños óptimos.

En otras palabras, aunque las funciones criterio pueden diferir en su forma y enfoque, pueden conducir a la misma conclusión en términos de selección del diseño óptimo. Esto implica que hay cierta flexibilidad en la elección de la función criterio, siempre y cuando el diseño seleccionado cumpla con los objetivos y requisitos específicos del experimento.

En resumen, el teorema de equivalencia demuestra que distintas funciones criterio pueden producir los mismos diseños óptimos. Esto permite cierta libertad al investigador para elegir

la función criterio que mejor se ajuste a sus necesidades y objetivos, siempre y cuando el diseño seleccionado cumpla con los requisitos deseados en términos de precisión y eficiencia.

LÓPEZ, V. I., & RAMOS, R. (2007). UNA INTRODUCCIÓN A LOS DISEÑOS ÓPTIMOS. REVISTA COLOMBIANA DE ESTADÍSTICA, 30(1), 37-51.

3.4. Teorema de equivalencia

Debido a que el problema de optimización que se intenta resolver en esta investigación es de difícil resolución se utilizarán teoremas de equivalencia en la práctica para, así, poder comprobar si un diseño es óptimo.

El Teorema de Equivalencia en el diseño de experimentos establece que distintas funciones criterio pueden dar lugar a los mismos diseños óptimos. En otras palabras, si dos o más funciones criterio conducen a la selección del mismo diseño óptimo, se dice que esas funciones criterio son equivalentes.

Este teorema es relevante porque proporciona flexibilidad al investigador para elegir la función criterio que mejor se adapte a sus necesidades y objetivos, sin comprometer la calidad del diseño resultante. Aunque las funciones criterio pueden diferir en su formulación matemática y enfoque, si producen el mismo diseño óptimo, se consideran equivalentes en términos de su capacidad para identificar el mejor diseño posible.

El Teorema de Equivalencia resalta que no existe una única función criterio "correcta" o "mejor" en todos los casos. Dependiendo de los objetivos del experimento, las limitaciones de recursos y otros factores, diferentes funciones criterio pueden proporcionar resultados similares en términos de selección de diseño óptimo. Esto permite cierta flexibilidad en la elección de la función criterio, siempre y cuando se cumplan los requisitos y objetivos específicos del experimento.

El primer teorema de equivalencia fue demostrado por Kiefer y Wolfowitz en el año 1960 y dice lo siguiente:

Sea $\chi \subseteq \mathbb{R}^k$ con m vectores linealmente independientes. Un diseño ξ con matriz de momentos $\mathrm{M}(\xi)$, definida positiva, es D-óptimo si y sólo si ξ es G-óptimo si y sólo si $f^T(x)M(\xi)^{-1}f(x) \leq m, \forall x \in \chi$ si y sólo si $\bar{d}\big(M(\xi)\big) = m.$

En caso de optimalidad, $f^T(x_i)M^{-1}f(x_i) = m, \quad \xi(x_i) \leq \frac{1}{m}, \quad \forall x_i \in Supp(\xi).$

Por lo popular de los criterios $\Phi_p(p \in [-\infty, 1])$, se enuncia el siguiente teorema de equivalencia, da condiciones necesarias y suficientes para garantizar que un diseño dado es Φ_p-óptimo.

Demostración:

Sean $\xi^*, \xi_1 \in \xi$ un diseño D-óptimo y otro diseño cualquiera, respectivamente. Si definimos el diseño ξ_2 como una combinación lineal convexa de los otros dos, es decir, $\xi_2 = (1-\alpha)\xi^* + \alpha\xi_1$.

Entonces,

$\mathrm{M}(\xi_2) = (1-\alpha)\mathrm{M}(\xi^*) + \alpha\mathrm{M}(\xi_1).$

Ademas, la función

$-\log\det[(1-\alpha)\mathrm{M}(\xi^*)+\alpha\mathrm{M}(\xi_1)]$,

es convexa y diferenciable, y tiene un mínimo en $\alpha=0$. Se sigue que,

$$0\leq\left(-\frac{\partial}{\partial\alpha}\log\det[(1-\alpha)\mathrm{M}(\xi^*)+\alpha\mathrm{M}(\xi_1)]\right)_{\alpha=0}$$

Tenemos que

$$\frac{\partial}{\partial\alpha}(-\log det\,[(1-\alpha)M(\xi^*)+\alpha M(\xi_1)])=Tr\left(M^{-1}(\xi_2)\frac{\partial}{\partial\alpha}M(\xi_2)\right)$$
$$=-Tr\left(M^{-1}(\xi_2)\big(M(\xi_1)-M(\xi^*)\big)\right)$$

Evaluando en $\alpha=0$, tenemos que

$-\mathrm{Tr}\left(\mathrm{M}^{-1}(\xi^*)\big(\mathrm{M}(\xi_1)-\mathrm{M}(\xi^*)\big)\right)=-\mathrm{Tr}(\mathrm{M}^{-1}(\xi^*)\mathrm{M}(\xi_1)-\mathrm{I_m})=-\left(\mathrm{Tr}\big(\mathrm{M}^{-1}(\xi^*)\mathrm{M}(\xi_1)\big)-\mathrm{Tr}(\mathrm{I_m})\right)=\mathrm{m}-\mathrm{Tr}\big(\mathrm{M}^{-1}(\xi^*)\mathrm{M}(\xi_1)\big)\geq 0,$

y esto implica que

$$\mathrm{Tr}\big(\mathrm{M}^{-1}(\xi^*)\mathrm{M}(\xi_1)\big)\leq\mathrm{m}$$

Sea ξ_x el diseño unipuntual soportado en el punto x. Entonces,

$\mathrm{Tr}\big(\mathrm{M}^{-1}(\xi^*)\mathrm{M}(\xi_1)\big)-\mathrm{m}=\mathrm{Tr}\big(\mathrm{M}^{-1}(\xi^*)\mathrm{M}(\xi_\mathrm{x})\big)-\mathrm{m}=\mathrm{Tr}\big(\mathrm{M}^{-1}(\xi^*)\mathrm{f}(\mathrm{x})\mathrm{f}^{\mathrm{t}}(\mathrm{x})\big)-\mathrm{m}=\mathrm{f}^{\mathrm{t}}(\mathrm{x})\mathrm{M}^{-1}(\xi^*)\mathrm{f}(\mathrm{x})-\mathrm{m}\leq 0,$

por lo que

$$\mathrm{d}(\mathrm{x},\xi^*)\leq\mathrm{m},\quad\forall\mathrm{x}\in\mathrm{X}$$

Por otro lado, para cualquier diseño ξ,

$\int_x d(x,\xi)\xi(dx)=\int_x f^t(x)M^{-1}(\xi)f(x)\xi(dx)=\int_x Tr\big(f^t(x)M^{-1}(\xi)f(x)\big)\xi(dx)=\int_x Tr\big(M^{-1}(\xi)f(x)f^t(x)\big)\xi(dx)=Tr\big(M^{-1}(\xi)\big)\int_x f(x)f^t(x)\xi(dx)=Tr\big(M^{-1}(\xi)M(\xi)\big)=Tr(I_m)=m.$

Se sigue que

$$\int_x d(x,\xi)\xi(dx)=m\leq\max_x d(x,\xi)\int_x\xi(dx)=\max_x d(x,\xi).$$

Hemos demostrado esto para cualquier diseño ξ, por lo que tenemos

$$\max d(x,\xi^*)\geq m,\forall x.$$

Además, hemos demostrado que

$$\mathrm{d}(\mathrm{x},\xi^*)\leq\mathrm{m},\forall\mathrm{x}.$$

Así que el máximo de todos ellos es menor o igual que m. Es obvio ver que el mínimo entre todos los máximos se alcanza en ξ^*, y su valor es m.

Supongamos ahora que ξ^* es un diseño G-óptimo pero no D-óptimo. Por lo tanto, para algún diseño ξ_1 tendríamos

$$\mathrm{Tr}\left(\mathrm{M}^{-1}(\xi^*)\mathrm{M}(\xi_1)\right) - \mathrm{m} > 0.$$

Además, $\mathrm{M}(\xi_1)$ se puede expresar como una combinación lineal convexa de $\mathrm{n} \leq \frac{\mathrm{m}(\mathrm{m}+1)}{2} + 1$ matrices de rango 1,

$$\mathrm{M}(\xi_1) = \sum_{\mathrm{i}=1}^{\mathrm{n}} \xi(\mathrm{x_i})\mathrm{f}(\mathrm{x_i})\mathrm{f}^{\mathrm{t}}(\mathrm{x_i}).$$

Por otro lado,

$$\begin{aligned}\mathrm{Tr}\left(\mathrm{M}^{-1}(\xi)\mathrm{M}(\xi)\right) &= \sum_{\mathrm{i}=1}^{\mathrm{n}} \xi(\mathrm{x_i})\mathrm{Tr}\left(\mathrm{M}^{-1}(\xi^*)\mathrm{f}(\mathrm{x_i})\mathrm{f}^{\mathrm{t}}(\mathrm{x_i})\right) \\ &= \sum_{\mathrm{i}=1}^{\mathrm{n}} \xi(\mathrm{x_i})\mathrm{f}^{\mathrm{t}}(\mathrm{x_i})\mathrm{M}^{-1}(\xi^*)\mathrm{f}(\mathrm{x_i}) = \sum_{\mathrm{i}=1}^{\mathrm{n}} \xi(\mathrm{x_i})\mathrm{d}(\mathrm{x_i}, \xi^*).\end{aligned}$$

Como ξ^* es G-óptimo,

$$\max d(x, \xi^*) = m, \forall x,$$

por lo que tenemos

$$\mathrm{d}(\mathrm{x}, \xi^*) \leq \mathrm{m}, \forall \mathrm{x}.$$

Lo que implica que

$$\sum_{i=1}^{n} \xi(x_i)d(x_i, \xi^*) - m \leq \sum_{i=1}^{n} \xi(x_i)m - m = m\sum_{i=1}^{n} \xi(x_i) - m = 0,$$

Llegando a una contradicción, ya que hemos probado que

$$\mathrm{Tr}\left(\mathrm{M}^{-1}(\xi^*)\mathrm{M}(\xi_1)\right) - \mathrm{m} > 0.$$

ARROYO, C. (2022). DISEÑO ÓPTIMO DE EXPERIMENTOS: APLICACIONES PARA LA ECUACIÓN DE ANTOINE, DISEÑOS D-AUMENTADOS Y PAQUETE OPTEDR.

4. Criterios característicos

4.1. Introducción y definiciones

Para la literatura, los criterios característicos mas conocidos y utilizados son el criterio D-óptimo, el cual trata de minimizar el volumen del elipsoide de confianza, relacionado con la estimación de máxima verosimilitud de los parámetros del modelo y el criterio A-óptimo, cuyo objetivo es minimizar la traza de la matriz de información, o lo que es lo mismo, minimizar la suma o promedio de las varianzas de los estimadores de los parámetros.

Por otra parte, también mencionar el criterio conocido como E-óptimo que consiste en maximizar el mínimo autovalor de la matriz de información.

FUNCINES SIMÉTRICAS ELEMENTALES

Primeramente vamos a definir lo que son las funciones simétricas elementales. Las funciones simétricas elementales son un tipo especial de funciones utilizadas en matemáticas, especialmente en el álgebra y la teoría de polinomios. Estas funciones están relacionadas con los coeficientes de un polinomio y expresan propiedades fundamentales de las raíces o los factores de dicho polinomio.

En particular, las funciones simétricas elementales se definen en términos de las raíces de un polinomio y se denominan "simétricas" porque son invariantes bajo permutaciones de las raíces. Esto significa que el valor de la función no cambia si se altera el orden de las raíces.

En este apartado va a ser muy importante el concepto de “polinomio característico” y por ello vamos a hablar de él a continuación.

El polinomio característico de una matriz es un concepto importante en álgebra lineal. Está asociado a una matriz cuadrada y es utilizado para determinar algunas de sus propiedades fundamentales. El polinomio característico se define como el determinante de la matriz resultante al restarle a la matriz original un múltiplo de la matriz identidad.

El polinomio característico es una función polinómica de λ, y sus raíces se denominan valores propios (o autovalores) de la matriz A. Estos valores propios son de gran importancia en el análisis de matrices, ya que proporcionan información sobre la estructura y comportamiento de la matriz.

Teniendo una matriz M que es simétrica y semidefinido positiva de orden m, su polinomio característico se puede escribir de la siguiente manera:

$$P_c(M) = |xI - M| = \varphi_0(M)x^m - \varphi_1(M)x^{m-1} + \varphi_2(M)x^{m-2} + \ldots + (-1)^m\varphi_m(M)$$

donde,

$$\varphi_k(M) = \sum_{i_1 < \ldots < i_k} \lambda_{i_1} \ldots \lambda_{i_k}, \quad 1 \leq k \leq m$$

$$\varphi_0(M) = 1$$

son las funciones simétricas en los valores propios de la matriz M, $\lambda_1, \ldots, \lambda_m$. El primer y el último coeficiente se tratan de

$$\varphi_1(M) = Tr(M)\ y\ \ \varphi_m(M) = det(M)$$

dando lugar, estos, a los criterios de optimización ya conocidos y mencionados anteriormente, A y D optimización respectivamente:

$$\Phi_A[M(\xi)] = tr[M^{-1}(\xi)];\quad \Phi_D[M(\xi)] = \log det[M^{-1}(\xi)].$$

El polinomio característico también tiene otras propiedades relevantes. Por ejemplo, el grado del polinomio característico es igual al tamaño de la matriz, es decir, n. Además, el polinomio característico permite calcular el determinante de la matriz A, ya que evaluando el polinomio en $\lambda = 0$ se obtiene det(A).

Se conocen como funciones simétricas elementales o coeficientes característicos a las funciones $\varphi_1, \ldots, \varphi_m$. Por otra parte, estas funciones están caracterizadas por ser funciones lineales.

En ciertas situaciones, puede ocurrir que la matriz en consideración no esté completamente definida y dependa de los valores de algún otro parámetro. En tales casos, es conveniente calcular las funciones simétricas elementales o coeficientes característicos de la siguiente manera:

$$\varphi_k(M) = \sum_{i_1 < \ldots < i_k} det\left[M_{i_1,\ldots,i_k}\right],\quad i \le k \le m$$

Este enfoque nos permite obtener las funciones simétricas elementales o coeficientes característicos correspondientes, teniendo en cuenta las características y la estructura de la matriz en cuestión, incluso cuando no está completamente definida inicialmente. Esta técnica puede ser útil en diversos campos de las matemáticas y la estadística, donde es necesario calcular estas funciones para comprender mejor las propiedades de las matrices y los sistemas de ecuaciones asociados.

4.1.1. Propiedades

En este contexto, es importante destacar que todas las matrices involucradas se considerarán simétricas y semidefinidas positivas. El enfoque del estudio se centra en las matrices de información, particularmente en aquellas que son regulares, ya que es crucial evitar casos que se salgan de lo común.

Por lo tanto, la condición mencionada al principio de esta sección se considera una condición fuerte que va más allá de los requisitos necesarios para el presente estudio, pero de todos modos no representa un obstáculo.

Lo que primeramente se va a definir es la matriz M, que se trata de una matriz diagonal con elementos $\lambda_1, \ldots, \lambda_m$. Esto es así debido a que el polinomio característico de M es invariante por cualquier cambio de base que se le haga de la forma $A^{-1}MA$. Se considera A una matriz no singular.

Por lo tanto, si diagonalizamos la matriz M, podemos observar que los elementos en la diagonal corresponden a los valores propios $\lambda_1, \ldots, \lambda_m$.

Además, es evidente que existe una relación entre la mayoría de las funciones criterio y la matriz de información del diseño. Sin embargo, la obtención de esta matriz puede ser un proceso computacionalmente costoso. Pero como solución a este problema aparece la formula ya que el determinante de $M^{-1}\left(\varphi_m(M^{-1})\right)$ es justamente el inverso del determinante de la matriz M. Esto se refiere a todos los coeficientes característicos.

Las funciones $\varphi_{m-k}(M)$ tendrán un cálculo fácil por medio de la aplicación de la desigualdad de Hadamard, propiedad muy importante que se aplica solo a matrices cuadradas y que expresa lo siguiente:

Si N es una matriz semidefinido positiva se verifica lo siguiente

$$\det \mathrm{N} \leq \prod_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{n}_{\mathrm{ii}}$$

La desigualdad de Hadamard indica que el valor absoluto del determinante de una matriz cuadrada es menor o igual al producto de las normas de sus columnas. En otras palabras, el determinante de una matriz no puede ser mayor que el producto de las normas de sus columnas.

Esta desigualdad indica que el determinante de una matriz está acotado por el producto de las normas de sus filas (o columnas). En otras palabras, cuanto más pequeñas sean las normas de las filas (o columnas) de la matriz, menor será el determinante.

La desigualdad de Hadamard es útil en diversas áreas de las matemáticas y la ciencia, como el análisis numérico, la teoría de la información, la teoría de la probabilidad y la teoría de códigos, entre otros. Se utiliza para establecer límites en el determinante de matrices y para analizar la estabilidad y la calidad de los algoritmos.

Por otro lado, se tiene que si m es par y k=m/2:

$$\frac{\varphi_k(M)}{\varphi_k(M^{-1})} = det(M)$$

Como parte de nuestro enfoque en el procedimiento de recurrencia, es interesante obtener una relación entre $\varphi_k(M)$ a partir de los resultados previos $\varphi_1(M), \ldots, \varphi_{k-1}(M)$. Estos resultados se centran en buscar la conexión o relación deseada.

Dicho de otra manera, estamos buscando una forma de relacionar los resultados anteriores para así poder establecer un patrón o una fórmula que nos permita obtener los resultados posteriores de manera recursiva.

4.1.2. Funciones características.

El propósito de las funciones características es generar nuevos criterios de optimización utilizando los coeficientes característicos. Es relevante utilizar la matriz de información inversa, que está relacionada con la matriz de covarianza de las estimaciones de los parámetros del modelo, como se mencionó anteriormente.

La función característica de una matriz se refiere a los coeficientes característicos de su inversa. Esto se representa de la siguiente manera:

$$\Phi_{Ch_k}(M) = \varphi_k(M^{-1})$$

Para cualquier función real v definida sobre un convexo se verifica:

1. Si v es concava, entonces es log-cóncava

2. Si v es log-convexa, entonces es convexa

Sean M y N matrices mxm definido positivas, y H una matriz sxm de rango s $(s \leq m)$. Entonces se verifica:

$$(HM^{-1}H^t)^{-1} + (HN^{-1}H^t)^{-1} \leq [H(M+N)^{-1}H^t]^{-1}$$

La función

$$\mathrm{M} \mapsto [\mathrm{M}^{-1}]^{-1}_{\mathrm{i}_1,\ldots,1_\mathrm{k}}$$

es cóncava, es decir, si N y M son matrices no singulares de rango m y $0 \leq \alpha \leq 1$ se cumple

$$[(\alpha M + (1-\alpha)N^{-1})^{-1}]^{-1}_{i_1,\ldots,i_k} \geq \alpha[M^{-1}]^{-1}_{i_1,\ldots,i_k} + (1-\alpha)[N^{-1}]^{-1}_{i_1,\ldots,i_k}$$

La función criterio $\Phi_{Ch_k}(M) = \varphi_k(M^{-1})$ es convexa, k=1,...,m

4.2. Continuidad y diferenciabilidad

La continuidad y la diferenciabilidad de $\varphi_k(M)$ se pueden deducir directamente de la definición en este caso. A partir de la definición de continuidad y diferenciabilidad podemos llegar a la conclusión de que en este contexto específico se trata de polinomios homogéneos en las componentes de M.

Lo anteriormente expuesto en este apartado implica que las funciones que describen la continuidad y diferenciabilidad dependan de manera lineal de las variables involucradas en la matriz M. El estudio presente se centra en el cálculo de su gradiente. Para ello se utiliza el siguiente resultado:

El gradiente de $\zeta_k(M) = \varphi_1(M^k)$ $(M\ simétrica)$ es

$$\nabla\, \zeta_k(M) = kM^{k-1}$$

El gradiente de los coeficientes característicos es

$$\nabla\varphi_\mathrm{k}(\mathrm{M}) = \sum_{\mathrm{j}=0,\ldots,\mathrm{k}-1} (-1)^\mathrm{j}\varphi_{\mathrm{k}-\mathrm{j}-1}(\mathrm{M})\mathrm{M}^\mathrm{j}$$

$\phi_k(M) = \varphi_k(M^{-1})$ es una función criterio diferenciable cuyo gradiente es

$$\nabla\phi_\mathrm{k}(\mathrm{M}) = \sum_{\mathrm{i}=\mathrm{k},\ldots,\mathrm{m}} \varphi_\mathrm{i}(\mathrm{M}^{-1})(-\mathrm{M})^{\mathrm{i}-\mathrm{k}-1} = -\sum_{\mathrm{i}=0,\ldots,\mathrm{k}-1} \varphi_\mathrm{i}(\mathrm{M}^{-1})(-\mathrm{M})^{\mathrm{i}-\mathrm{k}-1}$$

Una condición necesaria y suficiente para que un diseño ξ^* sea ϕ_k-óptimo es que se verifique lo siguiente

$$\min \sum_{i=k,\dots,m} (-1)^{-k+i-1}\varphi_i |M^{-1}(\xi^*)| f^t(x) M^{-k+i-1}(\xi^*) f(x)$$
$$= \sum_{i=k,\dots,m} (-1)^{-k+i-1}\varphi_i |M^{-1}(\xi^*)| \varphi_1 \left|M^{-k+i}(\xi^*)\right|$$

Es importante tener en cuenta que también se pueden utilizar funciones criterio más generales basadas en estos coeficientes. Por ejemplo, en lugar de limitarse a polinomios homogéneos en las componentes de la matriz, se pueden considerar otras funciones que involucren los coeficientes de manera más amplia.

Esto significa que no estamos limitados únicamente a polinomios homogéneos como criterio para evaluar o medir ciertas propiedades. Podemos explorar y utilizar funciones más generales que se basen en los coeficientes de la matriz, brindando así una mayor flexibilidad y capacidad de análisis en nuestro enfoque.

Para poner un ejemplo,

$$\phi_{p,k}(M) = \left[\frac{1}{\binom{m}{k}} \sum_{i_i<\dots<i_k} \left(\lambda_{i_i} \dots \lambda_{i_k}\right)^{-p/k}\right]^{1/p}$$

La división se utiliza para promediar, y la raíz k-ésima para volver a las unidades de medida originales. El resto se introduce como una generalización de los conocidos como criterios ϕ_p. En particular se podría expresar el criterio del determinante simplemente tomando p=1 y k=m, que es equivalente al definido, por ser log una función creciente. Opcionalmente se podrían haber definido tomando la raíz k-ésima global, en vez de hacerlo para cada sumando.

CEREZO, C., RAFAEL, J., & COSTA FUENTES, J. (2000). EPSF:ESTUDIO PREDICTIVO SOBRE FACTURACIÓN. UNA APLICACIÓN DEL ANÁLISIS DE SERIES TEMPORALES. SEIO, 14.

4.3. Interpretación geométrica

Es un resultado conocido que las longitudes de los ejes del elipsoide de confianza de las estimaciones de los parámetros del modelo son proporcionales a las raices cuadradas de los valores propios de la matriz de covarianzas. En este contexto, consideremos el caso de m=3 para facilitar la visualización de los conceptos.

El volumen del elipsoide de confianza es proporcional al producto de las longitudes de sus ejes. Por lo tanto, minimizar el producto de los valores propios implica minimizar el volumen del elipsoide (criterio de D-optimización). Asimismo, el área exterior de la caja que contiene al elipsoide se puede calcular como la suma de los productos de las longitudes de los ejes tomados de dos en dos. Así el criterio ϕ_2 minimiza un cierto promedio de las áreas de las proyecciones del elipsoide sobre los tres planos canónicos.

Estas proyecciones pueden interpretarse como las regiones de confianza de las estimaciones de diferentes pares de parámetros, y el diseño ϕ_2 óptimo como el que minimiza el promedio de las áreas de todas estas regiones de confianza. En general, el diseño ϕ_k-óptimo persigue minimizar el promedio de las áreas de las regiones de confianza de los estimadores de los distintos pares de parámetros del modelo, lo que a su vez está relacionado con el volumen del elipsoide de confianza y las proyecciones sobre los planos canónicos.

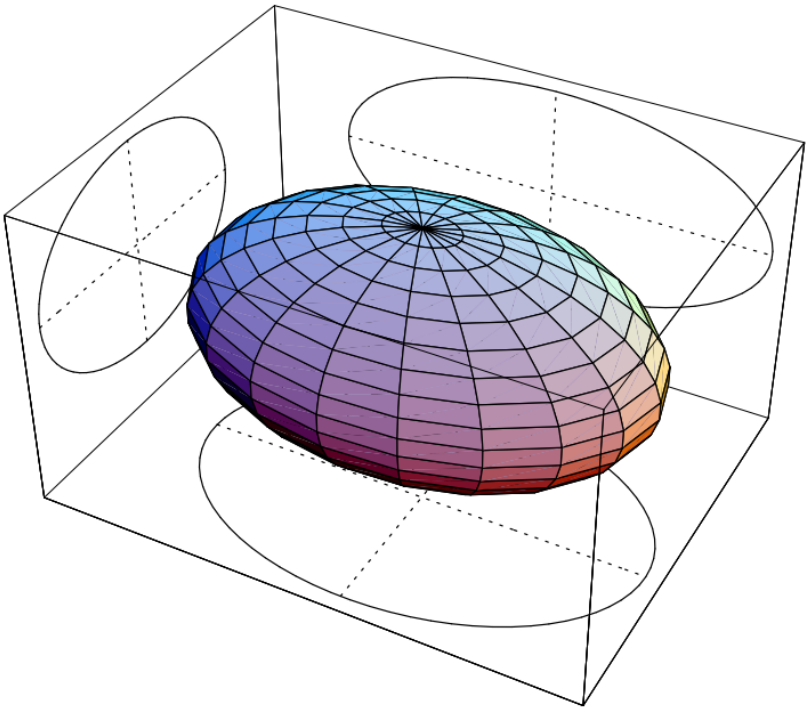


Figura 2: Interpretación geométrica del $\Phi \mathrm{Ch}_2$-óptimo para el modelo cuadrático.

CEREZO, C., RAFAEL, J., & COSTA FUENTES, J. (2000). EPSF:ESTUDIO PREDICTIVO SOBRE FACTURACIÓN. UNA APLICACIÓN DEL ANÁLISIS DE SERIES TEMPORALES. SEIO, 14.

5. Cálculo de óptimos característicos

5.1. Descripción del algoritmo

En el cálculo de los óptimos característicos, es importante simplificar los cálculos para facilitar el proceso. Por lo tanto, a partir de este punto, utilizaremos las funciones criterio que han sido definidas previamente

$$\Phi \mathrm{Ch}_{\mathrm{k}}(\mathrm{M}) = \varphi_{\mathrm{k}}(\mathrm{M}^{-1})$$

Al adoptar estas funciones criterio previamente establecidas, nos aseguramos de tener una base sólida y consistente para evaluar y determinar los óptimos característicos. Estas funciones criterio han sido cuidadosamente diseñadas y seleccionadas para reflejar las características y objetivos específicos del problema en cuestión. Al utilizarlas, simplificamos los cálculos y nos enfocamos en obtener resultados óptimos de manera más eficiente.

Los pasos a seguir del algoritmo son los siguientes:

I. Se elije un primer diseño ξ_0 en el que se verifique que el determinante de la matriz de información sea distinto de cero de la siguiente manera

$$\det \mathrm{M}(\xi_0) \neq 0$$

II. Iterativamente, se realiza la siguiente operación

$$\xi_{n+1} = (1 - \beta_n)\xi_n + \beta_n \xi_{x_n}$$

donde los parámetros β_n cumplen las siguientes características

$$\beta_n \in (0,1), \quad \sum_{n=0}^{\infty} \beta_n = \infty, \quad \lim_{n\to\infty} \beta_n = 0$$

el punto x_n es el punto en el que se alcanza el mínimo de la función

$$\mathrm{x} \in \mathrm{X} \to \mathrm{f}^{\mathrm{t}}(\mathrm{x}) \nabla \Phi[\mathrm{M}(\xi_{\mathrm{n}})] \mathrm{f}(\mathrm{x})$$

X es el espacio del diseño, $\mathrm{f}(\mathrm{x})$ el modelo y $\nabla \Phi[\mathrm{M}(\xi_{\mathrm{n}})]$ el gradiente de la función criterio Φ aplicado a la matriz de información del diseño ξ.

A continuación se va a realizar una explicación de lo que es el algortimo en si:

El algoritmo que se menciona ha sido desarrollado utilizando el programa de cálculo matemático llamado "Mathematica". Este programa incluye varias funciones independientes, como por ejemplo, la función que calcula la matriz de información de un diseño dado, el gradiente de una función criterio o el cálculo del mínimo de una función en un intervalo. Además, se incorpora un parámetro de precisión.

La función principal del algoritmo se encarga de buscar el óptimo y realiza refinamientos periódicos. Durante estos refinamientos, se eliminan los puntos cuyos pesos se consideran despreciables en ese momento. Luego, mediante un parámetro de proximidad, se distribuye el peso de estos puntos eliminados entre los puntos restantes, generando así un nuevo punto

que es la media de ambos o se reparte entre todos si no hay ninguno lo suficientemente cercano.

El usuario tiene la capacidad de ajustar varios parámetros, como el número de iteraciones en las que se realiza el refinamiento, el umbral a partir del cual un punto se considera de peso despreciable, la proximidad entre puntos y otros parámetros adicionales. Además, se puede establecer un parámetro de cercanía para indicar cuándo el diseño se considera lo "suficientemente cerca" del óptimo, lo que permite detener la ejecución del algoritmo, estableciendo así una condición de parada.

En resumen, el algoritmo desarrollado en el programa "Mathematica" utiliza varias funciones y parámetros configurables para buscar el óptimo de manera iterativa, realizando refinamientos y ajustes en el proceso hasta que se cumplan ciertas condiciones predefinidas por el usuario.

5.2. Modelo polinómico

El modelo polinómico es ampliamente utilizado en diversas disciplinas debido a su simplicidad y capacidad para aproximar funciones complejas de manera confiable. Este modelo ofrece una aproximación de buena calidad que resulta práctica en muchas situaciones. Los modelos polinómicos utilizan funciones polinómicas para aproximar datos y capturar relaciones no lineales. Se ajustan mediante regresión polinómica, minimizando la diferencia entre los valores observados y los pronósticos. Son útiles para representar diversas formas de datos, pero hay que tener cuidado con el sobreajuste al elegir el grado del polinomio.

En el caso del criterio D-óptimo, se puede obtener de manera explícita para el modelo polinómico, ya que su conjunto de soporte contiene las raíces o soluciones de los polinomios de Legendre. Además, existen otros resultados que demuestran que cuando el número de puntos en el soporte del diseño D-óptimo coincide con el número de parámetros del modelo, estos puntos tienen pesos idénticos. Estos resultados permiten calcular de manera precisa el diseño óptimo bajo este criterio.

En resumen, el modelo polinómico es muy utilizado debido a su capacidad para aproximar funciones complejas de manera sencilla. En el caso del criterio D-óptimo, se pueden obtener soluciones explícitas para el modelo polinómico, lo que facilita su cálculo y aplicación en la práctica.

Es fácil darse cuenta de que los $\Phi\mathrm{Ch}_k$-óptimos tienen soporte en tantos puntos como parámetros tiene el modelo, resultado ya conocido para k=1 y k=m (A y D-optimización respectivamente). Además, es fácil apreciar que los primeros $\Phi\mathrm{Ch}_k$-óptimos dan menos importancia a los puntos extremos del intervalo, importancia que va aumentando a medida que aumenta k hasta llegar al $\Phi\mathrm{Ch}_m$-óptimo (es decir, el D-óptimo), para el que los m puntos escogidos son igualmente importantes. Por lo demás, parece existir una “suave” transición entre el A y D-óptimos a través de los criterios intermedios.

A continuación vamos a presentar los datos de la aplicación del algoritmo para el cálculo de los óptimos característicos para los modelos de grado 2 a grado 6.

n	k	Óptimos para modelos polinómicos de grado n						
2	1	-1	0	1				
		0,25	0,5	0,25				
	2	-1	0	1				
		0,297	0,407	0,297				
	3	-1	0	1				
		0,333	0,333	0,333				
3	1	-1	-0,464	0,464	1			
		0,15	0,35	0,35	0,15			
	2	-1	-0,424	0,424	1			
		0,173	0,327	0,327	0,173			
	3	-1	-0,435	0,435	1			
		0,215	0,285	0,285	0,215			
	4	-1	-0,447	0,447	1			
		0,25	0,25	0,25	0,25			
4	1	-1	-0,667	0	0,667	1		
		0,105	0,25	0,29	0,25	0,105		
	2	-1	-0,643	0	0,643	1		
		0,116	0,256	0,256	0,256	0,116		
	3	-1	-0,633	0	0,633	1		
		0,139	0,232	0,257	0,232	0,139		
	4	-1	-0,643	0	0,643	1		
		0,17	0,216	0,228	0,216	0,17		
	5	-1	-0,655	0	0,655	1		
		0,2	0,2	0,2	0,2	0,2		
5	1	-1	-0,789	-0,291	0,291	0,789	1	
		0,08	0,187	0,232	0,232	0,187	0,08	
	2	-1	-0,767	-0,267	0,267	0,767	1	
		0,086	0,198	0,216	0,216	0,198	0,086	
	3	-1	-0,749	-0,28	0,28	0,749	1	
		0,099	0,195	0,206	0,206	0,195	0,099	
	4	-1	-0,747	-0,278	0,278	0,747	1	
		0,116	0,179	0,205	0,205	0,179	0,116	
	5	-1	-0,756	-0,281	0,281	0,756	1	
		0,141	0,173	0,186	0,186	0,173	0,141	
	6	-1	-0,765	-0,285	0,285	0,765	1	
		0,167	0,167	0,167	0,167	0,167	0,167	
6	1	-1	-0,853	-0,479	0	0,479	0,853	1
		0,065	0,148	0,185	0,205	0,185	0,148	0,065
	2	-1	-0,839	-0,451	0	0,451	0,839	1
		0,069	0,156	0,182	0,186	0,182	0,156	0,069
	3	-1	-0,824	-0,455	0	0,455	0,824	1
		0,076	0,162	0,166	0,191	0,166	0,162	0,076
	4	-1	-0,815	-0,461	0	0,461	0,815	1
		0,086	0,155	0,171	0,175	0,171	0,155	0,086
	5	-1	-0,817	-0,46	0	0,46	0,817	1
		0,1	0,145	0,168	0,173	0,168	0,145	0,1
	6	-1	-0,823	-0,464	0	0,464	0,823	1
		0,121	0,144	0,156	0,158	0,156	0,144	0,121
	7	-1	-0,83	-0,469	0	0,469	0,83	1
		0,143	0,143	0,143	0,143	0,143	0,143	0,143

Tabla 1: "Óptimos característicos para los primeros modelos polinómicos"

Por otro lado, también hay que tener en cuenta las eficiencias de los diseños que han sido calculados anteriormente respecto de las propias funciones criterio características. A continuación se presenta la tabla de las eficiencias mencionadas:

		Eficiencias						
n	k	1	2	3	4	5	6	7
2	1		98,5	94,5				
	2	98,8		98,9				
	3	96,2	98,9					
3	1		98,9	96,1	91,7			
	2	99,2		98,5	95			
	3	97,9	98,8		99			
	4	96,1	96	99				
4	1		99	97,4	94,6	90,7		
	2	99,3		99,2	96,9	93,2		
	3	98,5	99,4		99	96,3		
	4	97,7	98	99,1		99,1		
	5	96,6	96	96,8	99,2			
5	1		99,1	97,7	96,4	94	90,6	
	2	99,4		99,3	98,2	95,9	92,5	
	3	98,7	99,4		99,6	97,9	94,9	
	4	98,2	98,8	99,6		99,2	97	
	5	97,7	97,5	98,5	99,3		99,3	
	6	97	96,3	96,6	97,3	99,3		
6	1		99,2	98	96,8	95,8	93,8	90,7
	2	99,4		99,4	98,4	97,4	95,4	92,3
	3	98,8	99,5		99,6	98,8	97,1	94,2
	4	98,5	98,9	99,6		99,7	98,3	95,8
	5	98,2	98,4	99,1	99,7		99,4	97,6
	6	97,8	97,7	98,2	98,7	99,4		99,4
	7	97,3	96,7	96,7	97	97,8	99,4	

Tabla 2: "Tabla de eficiencias de los ΦCh_k-óptimos para los modelos polinómicos de grado n"

Como era de esperar, a medida que aumentamos el valor de k en cada modelo, la eficiencia del diseño óptimamente reduce. Sin embargo, es evidente que esta disminución es más pronunciada cuando k tiene pequeños valores en comparación con los casos en los que k se acerca al valor de m. Siguiendo este argumento, nos vamos a dar cuenta de un detalle, y es que resulta sorprendente que el mejor valor no lo obtenemos con k=m sino con valores k=m-1 o k=m-2.

En resumen, el objetivo del estudio es encontrar una combinación convexa adecuada de los criterios de A-optimización y D-optimización. Dado que el tratamiento de estos criterios no es numéricamente sencillo, se propone utilizar los criterios característicos intermedios como solución al problema de calcular los óptimos. Estos criterios ofrecen una solución rápida y eficiente, teniendo en cuenta lo observado anteriormente.

5.3. Otros modelos

Además del modelo polinómico, los criterios característicos también se utilizan en otros modelos, como el modelo exponencial y el modelo compartimental. Estos modelos serán explicados de manera clara y concisa a continuación.

Modelo exponencial.

En este modelo, la variable dependiente se representa como una suma de funciones exponenciales, donde cada una de estas funciones está ponderada por un parámetro lineal. Esto permite describir de manera más precisa fenómenos complejos que involucran múltiples procesos de crecimiento o decaimiento.

Se representa de la siguiente forma:

$$\eta(x, \theta) = \theta_1 + \theta_2 e^{x} + \theta_3 e^{-x}$$

en el intervalo [0,1].

Este modelo es comúnmente utilizado en diversas áreas, como la biología, la economía, la física, entre otras, para modelar fenómenos que exhiben un crecimiento o decaimiento exponencial, como el crecimiento de una población, la descomposición radioactiva o el crecimiento de una inversión financiera.

En la aplicación práctica del modelo exponencial, es importante ajustar los parámetros A y k a través de técnicas estadísticas, como el método de los mínimos cuadrados, para obtener una buena aproximación de los datos observados y realizar predicciones futuras.

A continuación, y de igual manera que con el modelo polinómico vamos a presentar su tabla de óptimos y su tabla de eficiencias.

k	óptimos		
1	0	0,498	1
	0,263	0,516	0,223
2	0	0,482	1
	0,382	0,354	0,266
3	0	0,498	1
	0,331	0,332	0,332

Tabla 3: "Tabla de óptimos característicos para el modelo exponencial"

k	Eficiencias		
1		89,3	86,84
2	92,12		98,08
3	93,27	98,75	

Tabla 4: "Tabla de eficienias para el modelo exponencial"

A la vista de los resultados expuestos es fácil declarar que, para este modelo, el ΦCh_3-óptimo resulta el mas eficiente de forma global. Cabe destacar que el ΦCh_2-óptimo se acerca mucho a este, es decir, ofrece también buenos resultados.

Modelo compartimental.

Los modelos compartimentales son herramientas utilizadas en diversas disciplinas, como la farmacología, la epidemiología y la ecología, para describir el comportamiento de sistemas que involucran la interacción de diferentes componentes o compartimentos. Estos modelos se basan en la idea de que el sistema en estudio puede ser dividido en compartimentos discretos, donde cada compartimento representa una entidad o una etapa específica del proceso.

En un modelo compartimental, los compartimentos se conectan mediante flujos, que representan las transferencias de masa, energía o información entre ellos. Estos flujos pueden ser uni o bidireccionales, dependiendo de la naturaleza del sistema y de las interacciones que se estén modelando.

El objetivo principal de los modelos compartimentales es comprender y predecir cómo cambian las cantidades o propiedades de interés en cada compartimento a lo largo del tiempo. Para lograr esto, se utilizan ecuaciones diferenciales que describen cómo evolucionan las variables de estado en función de las tasas de transferencia entre compartimentos.

En muchos casos, los modelos compartimentales se basan en principios de conservación de masa o energía, lo que implica que la suma de las cantidades en todos los compartimentos se mantiene constante. Esto permite representar fenómenos como la distribución de fármacos en el cuerpo humano, la propagación de enfermedades en una población o los flujos de nutrientes en un ecosistema.

Los modelos compartimentales pueden ser muy útiles para entender la dinámica de sistemas complejos y realizar predicciones sobre su comportamiento futuro. Sin embargo, es importante tener en cuenta que estos modelos son simplificaciones de la realidad y están sujetos a suposiciones y limitaciones inherentes. Por lo tanto, su validez y utilidad dependen en gran medida de la calidad de los datos utilizados para construirlos y de las hipótesis subyacentes en su formulación.

El hecho de elegir este modelo para definirlo y explicarlo se debe a varias razones, entre las que se destacan las siguientes:

- La necesidad de aplicar todo lo que ha sido explicado a lo largo del presente documento en situaciones reales, ya que este modelo es uno de los modelos que se utilizan con mayor frecuencia en la modelización de una gran diversidad de fenómenos reales.
- La popularidad que tienen en los ambientes anteriormente mencionados, pues a esto se debe el hecho de que exista gran cantidad de literatura acerca de este tema.
- Es un modelo no lineal, que apenas ha sido explicado con anterioridad en el presente documento.

Un modelo compartimental se puede definir así:

Un modelo compartimental es una representación simplificada de un sistema físico o biológico en el que se divide en un número finito de componentes o compartimentos. Estos compartimentos interactúan entre sí a través de transferencias de materia, ya sea

internamente o con el entorno externo. Dependiendo de si hay o no intercambio de materia con el exterior, el modelo puede ser abierto o cerrado respectivamente.

Cada compartimento se considera homogéneo, lo que significa que se asume que las propiedades dentro de cada compartimento son uniformes y se pueden describir de manera similar. Cada compartimento representa una situación o estado específico de la materia dentro del sistema en estudio.

El modelo compartimental general se construye mediante la combinación de funciones exponenciales, lo que implica que se utilizan ecuaciones que involucran exponenciales para describir las tasas de cambio de las cantidades en cada compartimento a lo largo del tiempo. Estas ecuaciones pueden variar según las características y objetivos específicos del sistema que se está modelando.

El modelo se representa de la siguiente manera:

$$\eta(x,\theta) = \theta_3\left[e^{-\theta_2 x} - e^{\theta_1 x}\right];\ \theta_2 > \theta_1 > 0, \theta_3 > 0$$

A continuación se presentan la tabla de óptimos y su correspondiente tabla de eficiencias:

k	óptimos		
1	0,185	1,264	23,426
	0,303	0,581	0,115
2	0,225	1,204	23,1
	0,371	0,527	0,102
3	0,229	1,387	18,405
	0,331	0,336	0,333

Tabla 5: "Tabla de óptimos característicos para el modelo compartimntal"

k	Eficiencias		
1		95,59	71,39
2	93,29		80,12
3	78,6	79,07	

Tabla 6: "Tabla de eficiencias para el modelo compartimental"

En este caso, al analizar el modelo compartimental, se observa que el ΦCh_2-óptimo ofrece mejores resultados en términos de eficiencia, llegando a alcanzar un porcentaje de mas del 80%.

6. Conclusiones

i. En relación al desarrollo de algoritmos para buscar diseños óptimos en conjuntos de puntos discretos, se puede concluir que utilizar estos algoritmos es más eficiente cuando el número de puntos es menor. Esto se debe a que el tiempo de cálculo necesario será reducido. Además, si los resultados deseados no requieren una gran precisión, también es posible aplicar estos algoritmos a conjuntos de puntos continuos. En resumen, al utilizar algoritmos para buscar diseños óptimos, es importante considerar el tamaño del conjunto de puntos y la precisión necesaria en los resultados. Si el número de puntos es pequeño y no se requiere una gran precisión, se pueden obtener resultados de manera más eficiente. Sin embargo, si se necesita una alta precisión o el conjunto de puntos es muy grande, puede ser necesario utilizar métodos más sofisticados o realizar aproximaciones para agilizar el cálculo.

ii. Introducción de una nueva clase de funciones criterio, los que ya se conocen como Criterios Característicos. Este tipo nuevo de criterios están basados en los coeficientes del polinomio característico de la matriz de información. Por este motivo, se tratan de generalizaciones de los criterios A-óptimo y D-óptimo. Como consecuencia, los nuevos criterios llevarán a definir un criterio mas general englobando, tanto a los nuevos Criterios Característicos como a los ya conocidos como criterios Φ_p .

iii. Los nuevos Criterios Característicos exhiben una variedad de propiedades notables. Entre ellas, se encuentra la propiedad circular que se cumple en las funciones simétricas en relación a los valores propios. Además, se observa la generación iterativa de cada función a partir de un subíndice inferior, así como la relación entre estas funciones y las características. Otras propiedades importantes de los criterios generados incluyen la convexidad y la diferenciabilidad. Estas propiedades resaltan la utilidad y versatilidad de los Criterios Característicos, ya que permiten un análisis riguroso y eficiente en la búsqueda de diseños óptimos.

iv. Las nuevas funciones criterio han sido diseñadas de manera que se favorezca la fácil aplicabilidad de algoritmos tradicionales en la búsqueda de diseños óptimos. Estas funciones presentan características específicas que permiten que los algoritmos tradicionales se puedan utilizar de manera sencilla y eficiente para alcanzar los diseños óptimos deseados. Esta facilidad de aplicación de los algoritmos tradicionales es una ventaja significativa, ya que aprovecha la experiencia y el conocimiento existente en el campo de la optimización para obtener resultados precisos y confiables en el diseño de experimentos.

v. En cuanto a la utilización de los nuevos criterios característicos para el cálculo de diseños óptimos característicos para modelos polinómicos de diferentes grados, cabe destacar que se comprueba que resultan altamente eficientes al ser comparados, tanto entre sí como comparados con los criterios más extremos, A-optimización y D-optimización.

vi. Obtención de otro tipo de modelos. El cálculo de óptimos característicos no se limita únicamente al modelo polinómico, sino que también puede aplicarse para obtener

óptimos en otros tipos de modelos. Por ejemplo, se puede utilizar en el modelo exponencial y el modelo compartimental, los cuales son ampliamente utilizados en ciencias experimentales. Estos modelos, al igual que el polinómico, pueden beneficiarse de los criterios característicos para encontrar diseños óptimos que maximicen la información obtenida y minimicen la incertidumbre en las estimaciones de los parámetros. De esta manera, se amplía el alcance y la aplicabilidad de los criterios característicos en la búsqueda de diseños óptimos en diversos contextos científicos y experimentales.

7. Bibliografia

ARROYO, C. (2022). DISEÑO ÓPTIMO DE EXPERIMENTOS: APLICACIONES PARA LA ECUACIÓN DE ANTOINE, DISEÑOS D-AUMENTADOS Y PAQUETE OPTEDR.

CANAVOS, G. C., & MEDAL, E. G. U. (1987). PROBABILIDAD Y ESTADÍSTICA (P. 651). MÉXICO: MCGRAW HILL.

CARMONA, F. (2005). MODELOS LINEALES. *PUB. UNIV. DE BARCELONA, BARCELONA*.

CEREZO, C., RAFAEL, J., & COSTA FUENTES, J. (2000). EPSF: ESTUDIO PREDICTIVO SOBRE FACTURACIÓN. UNA APLICACIÓN DEL ANÁLISIS DE SERIES TEMPORALES. SEIO, 14.

DÍAZ, A. (2009). *DISEÑO ESTADÍSTICO DE EXPERIMENTOS 2A ED*. UNIVERSIDAD DE ANTIOQUIA.

GARCÍA-CAMACHA GUTIÉRREZ, I. (2018). DISEÑO ÓPTIMO DE EXPERIMENTOS PARA MODELOS DE MEZCLAS APLICADOS EN LA INGENIERÍA Y EN LAS CIENCIAS EXPERIMENTALES.

ISGRO, C. A. (1991). ESTIMACION DE REGRESION A TROZOS (DOCTORAL DISSERTATION).

LÓPEZ, V. I., & RAMOS, R. (2007). UNA INTRODUCCIÓN A LOS DISEÑOS ÓPTIMOS. REVISTA COLOMBIANA DE ESTADÍSTICA, 30(1), 37-51.

MOSQUERA-BENÍTEZ, J. C., & LÓPEZ-RÍOS, V. I. (2019). EFECTO DE DISTRIBUCIONES A PRIORI EN LOS DISEÑOS D-ÓPTIMOS BAYESIANOS PARA UN MODELO NO LINEAL CORRELACIONADO. CIENCIA EN DESARROLLO, 10(2), 165-170.

NARANJO-PALACIOS, F (2019). DISEÑOS EXPERIMENTALES DE TAGUCHI FRACCIONADOS.

NARANJO-PALACIOS, F., RIOS-LIRA, A. J., PANTOJA-PACHECO, Y. V., & TAPIA-ESQUIVIAS, M. (2020). DISEÑOS ORTOGONALES DE TAGUCHI FRACCIONADOS. INGENIERÍA, INVESTIGACIÓN Y TECNOLOGÍA, 21(2).

NEGRON PEREZ, G. A. (2014). DISEÑOS OPTIMOS PARA EL MODELO DE MICHAELIS-MENTE.

PAGURA, J. A., & PUIGSUBIRÁ, C. (2008). DISEÑO ROBUSTO DE PARAMETROS. ALTERNATIVAS BASADAS EN DISEÑO ÓPTIMO.

RODRÍGUEZ HERNÁNDEZ, M. D. L. M. (2016). DISEÑO ÓPTIMO DE EXPERIMENTOS PARA MODELOS DE ADAIR Y MEZCLA DE DISTRIBUCIONES.

RODRÍGUEZ DÍAZ, J. M (2001). CRITERIOS CARACTERÍSTICOS EN DISEÑO ÓPTIMO DE EXPERIMENTOS (VOL. 67). UNIVERSIDAD DE SALAMANCA.

RODRÍGUEZ-JAUME, M. J., & MORA CATALÁ, R. (2001). ANÁLISIS DE REGRESIÓN MÚLTIPLE.

SEIER, E. (1990). MATRIZ DE PROYECCIÓN Y DISEÑOS ÓPTIMOS PARA REGRESIÓN. PRO MATHEMATICA, 4(7-8), 31-69.

VNiVERSiDAD D SALAMANCA
CAMPUS DE EXCELENCIA INTERNACIONAL

Facvltad D Ciencias
VNiVERSiDAD D SALAMANCA

Certificado de los tutores TFG Grado en Estadística

D. Juan Manuel Rodríguez Díaz, profesor/a del Departamento de Estadística de la Universidad de Salamanca,

HACEN CONSTAR:

Que el trabajo titulado "Criterios de Optimización Característicos", que se presenta, ha sido realizado por D. Julio González Almeida, con DNI 45690366-S y constituye la memoria del trabajo realizado para la superación de la asignatura Trabajo de Fin de Grado en Estadística en esta Universidad.

Salamanca, a fecha de firma electrónica.

RODRIGUEZ DIAZ JUAN MANUEL - 07960523Q

Firmado digitalmente por RODRIGUEZ DIAZ JUAN MANUEL - 07960523Q
Fecha: 2023.07.03 10:42:06 +02'00'

Fdo.: Julio González Almeida

Fdo.: Juan Manuel Rodríguez Díaz