

# DISEÑO DE EXPERIMENTOS EN METROLOGÍA

Román de la Vara Salazar  
Centro de Investigación en Matemáticas  
Callejón de Jalisco, S/N, La Valenciana, Guanajuato, Gto.  
Tel. (473)7327155, Fax: (473) 7325749, Email: delavara@cimat.mx

**Resumen:** El Diseño de Experimentos es una de las herramientas estadísticas con mayor aplicación en las diferentes áreas de conocimiento. En este trabajo se discute la relevancia del diseño estadístico de experimentos en la Metrología. Para ello primero se da una panorámica general sobre el diseño de experimentos y los diferentes problemas que puede ayudar a resolver. Se revisan los pasos para la planeación y ejecución de un experimento, y finalmente se presenta una aplicación en la estimación de incertidumbre.

## INTRODUCCIÓN

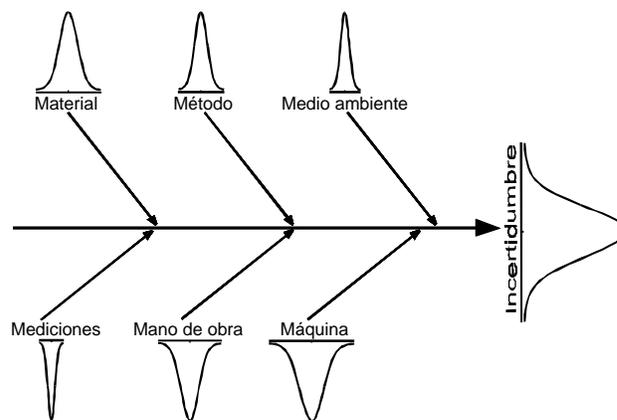
El origen del Diseño Estadístico de Experimentos se remonta a principios del siglo pasado cuando Ronald Fisher era el responsable de estadística y del análisis de datos en un centro de investigación agrícola cerca de Londres, Inglaterra. El problema al que se enfrentó Fisher fue que no era posible dar respuesta a las preguntas planteadas por investigadores a partir de los datos experimentales que obtenían. La solución fue que antes de realizar las corridas experimentales, las cuales podían llevar meses, Fisher los asesoraba acerca de cómo debían realizar tales pruebas. De esta manera surgió el diseño de experimentos.

Los primeros diseños experimentales fueron de tipo comparativo, pero poco a poco se fueron agregando otros arreglos como los diseños factoriales, y más recientemente (1980's) los diseños robustos o de Taguchi [5], hasta conformar en la actualidad distintas áreas de investigación y de aplicación dentro de lo que se conoce como diseño de experimentos; tan es así que difícilmente un profesional de este campo llega a conocer siquiera la mayoría de los diseños experimentales que se han propuesto. El surgimiento de tantos diseños experimentales radica principalmente en la amplia gama de procesos que existen para fabricar todo tipo de productos.

Cada vez son más las empresas y centros de investigación y desarrollo que aplican el diseño de experimentos de manera cotidiana, tanto para mejorar el desempeño de procesos ya establecidos como para hacer innovación y diseño de nuevos productos. Esto se ha hecho necesario para mantener la empresa competitiva en un mercado global, al poder reducir sus costos de producción

mejorando al mismo tiempo la calidad de sus productos. De este esfuerzo no son la excepción los laboratorios de metrología.

La medición se puede visualizar como el resultado de un proceso, que es precisamente el proceso de medición. Y como en cualquier proceso su resultado está sujeto a la influencia de factores que tradicionalmente en la industria se han clasificado en las seis ramas: Material, Mano de obra (operadores), Máquinas (equipos), Medio ambiente, Método y Mediciones (Figura 1). Prácticamente cualquier factor que afecta el resultado del proceso se puede ubicar en alguna de estas ramas y en cada una de estas ramas se encuentran varios factores específicos que pueden ser los causantes de la calidad o variabilidad del resultado del proceso.



**Fig. 1** Fuentes de incertidumbre en un proceso de medición.

En este trabajo se discute la relevancia del diseño estadístico de experimentos en la solución de problemas de metrología. Para ello comenzamos definiendo lo que es un diseño de experimentos, damos una clasificación de los esquemas experimentales de mayor aplicación de acuerdo al tipo de problema que resuelven y se revisan los pasos para la planeación y ejecución de un experimento. Finalmente se muestra un ejemplo en la estimación de incertidumbre.

## **ALGUNAS DEFINICIONES Y CLASIFICACIÓN DE LOS DISEÑOS EXPERIMENTALES**

Antes de dar una clasificación de los diseños experimentales revisemos algunas definiciones útiles.

**Diseño de Experimentos (DE).** El DE consiste en planear un conjunto de pruebas o corridas experimentales de tal manera que al analizar los datos que de ellas se obtengan se puedan contestar las preguntas planteadas. Esto no implica que de un experimento se puedan obtener todas las preguntas y todas las respuestas, sino que se pueden necesitar varios experimentos secuenciales para llegar a un resultado satisfactorio [9].

**Variables, Factores y Niveles.** En cualquier estudio experimental intervienen varios tipos de variables. La variable de respuesta tiene que ver con el resultado del experimento; es la variable que se está estudiando como función de los llamados factores. Los factores son las variables cuyo efecto sobre la respuesta interesa estudiar. Para ello en cada factor se eligen cuando menos dos niveles de prueba para poder estudiar su efecto. Es precisamente las diferentes combinaciones de niveles de los factores lo que constituyen el llamado arreglo experimental o matriz de diseño.

Cabe señalar que existen factores económicos de controlar durante los experimentos, que muchas veces algunos de ellos ya se “controlan” para obtener un buen resultado. La palabra controlar aquí es en el sentido de poder fijarlo en algún valor específico. También existen factores que no son económicos de controlar y que pueden influir en el proceso de medición. Si estos se pudieran controlar al menos para fines experimentales se pueden incluir en el estudio como factores de ruido. La idea con los factores de ruido es encontrar la combinación de los factores de control donde el efecto de los primeros sobre el resultado del proceso sea mínimo. Esta es la idea del llamado diseño robusto.

**El concepto de bloque.** Un bloque es un conjunto de resultados experimentales que se esperan más homogéneos que el conjunto total de resultados. Al considerar el efecto del bloque se logra una mayor precisión del experimento. Por ejemplo, si durante el experimento es necesario utilizar material de dos lotes diferentes, cada lote constituye un bloque. Las corridas experimentales se reparten en forma balanceada entre los bloques y al final se podrá separar el efecto debido al factor bloque.

### **Clasificación de los diseños experimentales.**

El objetivo del experimento se ha utilizado tradicionalmente como un criterio general de clasificación de los diseños experimentales. Otros aspectos como son el número de niveles, las restricciones existentes y el número de corridas se utilizan para hacer una subclasificación más fina. De acuerdo a su objetivo se clasifican como:

1. **Diseños para comparar tratamientos.** En este grupo se encuentran el diseño completamente al azar, el diseño en bloques completos al azar, diseño en cuadro latino, diseños en bloques incompletos, etc. Con estos diseños se pueden comparar diferentes métodos de medición, varios instrumentos, varios operadores, diferentes laboratorios, etc. Estos diseños no consideran la posible interacción entre los factores estudiados.
2. **Diseños para estudiar el efecto de varios factores sobre una o varias respuestas.** En esta rama se encuentran los diseños factoriales completos y fraccionados con dos o más niveles en cada factor. Una característica distintiva de estos arreglos es que permiten el estudio de efectos de interacción entre los factores. Los diseños factoriales tienen amplia aplicación en metrología, por ejemplo en la determinación de incertidumbre tipo A y en los estudios de repetibilidad y reproducibilidad (R&R) [3][2].
3. **Diseños para la optimización de procesos.** En esta familia entran otra vez los diseños factoriales y algunas generalizaciones de éstos como son el diseño de Box-Behnken y el diseño central compuesto. Estos diseños tienen al menos tres niveles en cada factor lo que permite investigar efectos de tipo cuadrático, además de interacciones y efectos lineales [9]. Con estos diseños se puede optimizar un proceso de medición, encontrando las condiciones en las que arroja los mejores resultados.
4. **Diseños para la optimización de mezclas.** La característica distintiva de los procesos de mezclas es que su desempeño depende de la

proporción con la que participan los componentes de la mezcla, no de la cantidad de mezcla, y la suma de todos los componentes en cada punto de prueba suman la unidad. Entre diseños de esta familia se encuentran los diseños de lattice-simplex, simplex con centroide, y diseños con restricciones [1].

5. **Diseño robusto.** Estos diseños tienen como uno de sus objetivos el hacer el producto o proceso insensible a factores no controlados o de ruido. El diseño con arreglo interno y externo fue propuesto por Taguchi [5][8] con este objetivo en mente, pero también se puede hacer diseño robusto utilizando un solo arreglo factorial enfocándose al análisis de las interacciones entre los factores de control y de ruido [6]. Por ejemplo, en un laboratorio pudieran existir condiciones para mantener un material o instrumento patrón en las cuales algunos factores que no se pueden controlar tienen un efecto mínimo sobre el patrón. Con este tipo de experimentos se pueden investigar cuáles son esas condiciones mejores de conservación.

Por supuesto que en cada familia solo mencionamos los diseños más representativos, y habría algunos diseños importantes no mencionados en la clasificación anterior como son los diseños anidados, cíclicos y de parcelas divididas (split plot). Los diseños anidados son bastante utilizados en los estudios R&R. Por su parte, los diseños factoriales podrían muy bien ubicarse en cualquiera de los objetivos anteriores dado que su versatilidad los hace sumamente aplicables. Por cierto que el segundo objetivo es de hecho el objetivo general que tiene cualquier diseño experimental: En cualquier DE se quiere estudiar el efecto de uno o varios factores sobre una o varias respuestas.

## **PASOS EN LA PLANEACIÓN Y EJECUCIÓN DE UN EXPERIMENTO**

Muchas veces, de manera simplista, se le llama “el diseño” al arreglo experimental, pero la realidad es que el diseño de experimentos es mucho más que un listado de corridas experimentales. Quizás la etapa más importante es la planeación del estudio donde entra tanto el pensamiento estadístico como el conocimiento físico sobre el proceso a estudiar. Debe anticiparse desde la planeación cualquier problema que pueda surgir. Los pasos son:

1. Ubicar un proceso de metrología susceptible de mejorarse con diseño de experimentos. Tener bien claro cuál es el objetivo del experimento, cuál el impacto esperado en caso de tener éxito.
2. Elegir la variable de respuesta (la medición de interés) y verificar que se mide de manera confiable para fines del experimento.
3. Determinar cuáles factores deben controlarse o investigarse de acuerdo a la supuesta influencia que tienen sobre la variable de respuesta de interés.
4. Seleccionar el experimento adecuado a los factores que se tienen y al objetivo del experimento. En este paso deben tomarse en cuenta las restricciones existentes, los costos, el número de repeticiones, el tipo de efectos esperados, el tiempo y recursos en general que se invertirán. Uno de los aspectos más críticos es la selección de los niveles de prueba en cada factor.
5. Realizar las corridas experimentales aleatorizando de acuerdo al plan y manteniendo en el transcurso lo más fijo posible los factores no incluidos en el estudio. Se recomienda que la persona interesada siempre este presente.
6. Determinar el modelo de análisis de varianza (ANOVA) o el modelo que mejor describa el comportamiento de los datos.
7. Verificar que los supuestos del modelo se cumplen. Diagnosticar la calidad del modelo.
8. Utilizar el modelo para los fines del estudio. Por ejemplo, si se trataba de estimar componentes de incertidumbre, proceder a hacerlo, o si el objetivo era maximizar la respuesta encontrar el punto donde eso ocurre.

## **EJEMPLO**

De acuerdo a la Guía para la Evaluación de la Incertidumbre [4], se tienen incertidumbre tipo A y tipo B. La incertidumbre tipo A se calcula con medios estadísticos y la tipo B por cualquier otro medio. Uno de los medios para calcular incertidumbre tipo B es la especificación dada por el fabricante del equipo. Sin embargo, en esto muchas veces no se toma en cuenta el hecho de que dicha especificación describe a toda la población de instrumentos del fabricante y es posible que el nuestro se encuentre en una de las colas de dicha población o en el centro de la misma, en cuyo caso

su incertidumbre sería un poco mayor o menor que la especificación dada. El diseño de experimentos puede ayudar a resolver este problema, evaluando experimentalmente en las condiciones del laboratorio el efecto de los factores de influencia.

Consideremos un experimento [7] donde el objetivo es estudiar la relación entre la frecuencia de oscilación de un reloj de cuarzo patrón y las condiciones de humedad y temperatura. En este caso el instrumento ya cuenta con un dispositivo para minimizar los cambios de temperatura, dado que los fabricantes conocen su impacto en la frecuencia de oscilación. Los factores seleccionados son temperatura (T) y humedad (H) y sus niveles de prueba se eligen de acuerdo a las condiciones del laboratorio; en este caso los niveles de temperatura son (22°C, 24 °C) y para la humedad (20%, 50%). La variable de respuesta es la frecuencia de oscilación (Y). El diseño experimental seleccionado es un factorial completo 2<sup>2</sup> con punto central y dos réplicas, lo que da un total de 10 corridas que se muestran en la Tabla 1.

Bloque	Temp. (°C)	Hum. (%)	Frecuencia (Hz)
1	24 (+1)	20 (-1)	9 999 999.9697
1	22 (-1)	50 (+1)	9 999999.9678
1	23 (0)	35 (0)	9 999999.9692
1	24 (+1)	50 (+1)	9 999999.9690
1	22 (-1)	20 (-1)	9 999999.9706
2	22 (-1)	50 (+1)	9 999999.9704
2	24 (+1)	50 (+1)	9 999999.9702
2	23 (0)	35 (0)	9 999999.9704
2	24 (+1)	20 (-1)	9 999999.9706
2	22 (-1)	20 (-1)	9 999999.9717

**Tabla 1.** Experimento 2<sup>2</sup> más centro en dos bloques y frecuencias observadas

En este caso cada réplica constituye un bloque, dado que se hicieron en días distintos, y es sabido que el oscilador de cuarzo presenta envejecimiento que ocasiona que su frecuencia derive en niveles conocidos. Por conveniencia de cómputo, se realiza el análisis de varianza para la respuesta transformada  $f_t = (\text{frecuencia} - 9999999.9)10^4$  y los resultados del análisis de varianza (ANOVA) se muestran en la Tabla 2, donde FV= Fuentes de variación, SC= Sumas de cuadrados, GL=Grados de libertad, CM= Cuadrados medios, F<sub>0</sub>= Estadístico F y P-valor o significancia estadística del término correspondiente. El P-valor es la probabilidad de

haber observado el efecto que ocurrió, suponiendo que dicho efecto poblacional fuera nulo. De manera que si el P-valor es pequeño comparado con  $\alpha$ , se concluye que el efecto o término correspondiente es estadísticamente significativo, es decir, tiene influencia sobre la variable de respuesta. El valor  $\alpha$  es la probabilidad de concluir que un efecto es significativo cuando en realidad es nulo, riesgo que se controla directamente por el experimentador al asignarle desde antes del experimento una magnitud pequeña, típicamente de 0.05.

F.V.	S.C.	G.L.	C.M.	F <sub>0</sub>	P-valor
Temp.	12.5	1	12.5	0.63	0.4638
Hum.	338.0	1	338.0	17.00	0.0091
TxH	112.5	1	112.5	5.66	0.0633
Bloque	490.0	1	490.0	24.65	0.0042
Error	99.4	5	19.88		
Total	1052.4	9			

**Tabla 2.** ANOVA

Se observa en la columna del P-valor que la temperatura no fue significativa (P-valor >  $\alpha=0.05$ ), mientras que la humedad y el efecto de bloque si lo fueron. La interacción temperatura-humedad no es significativa con  $\alpha=0.05$ , pero su efecto no es del todo aleatorio dado su P-valor cercano a 0.05; se esperaría que con más réplicas este efecto resultara significativo.

El modelo de ANOVA solo con los términos significativos tiene la forma

$$Y_{ijk} = m + H_i + (TH)_{ij} + Bloque_k + e_{ijk},$$

donde  $i, j, k = 1, 2$ . De acuerdo a este modelo, la variación total observada en el estudio se puede descomponer según el modelo de componentes de varianza

$$s_{total}^2 = s_H^2 + s_{TH}^2 + s_{Bloque}^2 + s_{Error}^2$$

Los componentes de varianza se pueden estimar a partir de los cuadrados medios calculados de la tabla de ANOVA, mediante las fórmulas [6]

$$s_{Error}^2 = CM_{Error} ; s_H^2 = \frac{CM_H - CM_{TH}}{2 \times 2}$$

$$s_{TH}^2 = \frac{CM_{TH} - CM_{Error}}{2} ; s_{Bloque}^2 = \frac{CM_{Bloque} - CM_{Error}}{2 \times 2}$$

En nuestro ejemplo tenemos, respectivamente,

$$s_{Error}^2 = CM_{Error} = 19.88; s_H^2 = \frac{338-112.5}{2 \times 2} = 56.3$$

$$s_{TH}^2 = \frac{112.5-19.8}{2} = 46.3; s_B^2 = \frac{490-19.8}{2 \times 2} = 117.5.$$

Pero como estos resultados son para la respuesta transformada  $ft=(frecuencia - 9999999.9)10^4$ , habría que multiplicarlos por  $10^{-8}$  para obtener el resultado en la escala de las frecuencias originales. Esto es porque, despejando, se tiene que la frecuencia  $= (ft)10^{-4} + 9999999.9$  y al momento de calcular la varianza en ambos lados resulta que  $Var(frecuencia) = 10^{-8} Var(ft)$ . Así, por ejemplo, el componente de varianza en la escala original para la humedad es de magnitud  $s_H^2 = 5.6 \times 10^{-7}$ . De aquí que la incertidumbre experimental debida este término está dada por  $s_H = 7.5 \times 10^{-4}$ . Esta incertidumbre observada se compara con la declarada por el fabricante del equipo.

## CONCLUSIONES

En general el diseño de experimentos puede contribuir bastante a mejorar la práctica metrológica, ayudando en la planeación y obtención de datos más informativos en los diferentes tipos de experimentos que se realizan en este campo. La aplicación de principios tales como aleatorización, bloqueo y repetición, así como el análisis estadístico correcto resultan de suma importancia. En particular en el estudio presentado se muestra cómo evaluar experimentalmente la incertidumbre tipo B dada por el fabricante de un equipo para verificar su magnitud en las condiciones del propio laboratorio. Este tipo de estudios podrían llevar a mejoras tanto de los equipos como de las instalaciones del laboratorio, buscando tener un menor impacto de las fuentes de incertidumbre detectadas como las más importantes.

## REFERENCIAS

- [1] Cornell, J. A., Experiments with Mixtures, Wiley, New York, 1990.
- [2] Carroll Croakin, M. y Tobias, P., Engineering Statistics Handbook. <http://www.itl.nist.gov/div898/handbook>.
- [3] Carroll Croakin, M., Statistics and Measurements, Journal of Research of the NIST 106, 2001, pp. 279-292.
- [4] ISO TAG4 WG3: Guide to the Expression of Uncertainty in Measurement, 1992.
- [5] Kacker, R. N., Off-Line Quality Control, Parameter Design and the Taguchi Method, Journal of Quality Technology 17, 1985, pp. 176-188
- [6] Montgomery, D. C., Design and Analysis of Experiments, Wiley, New York, 2001.
- [7] Montoya Vargas, A., Diseño de Experimentos en la Metrología, Taller de Metrología, CIMAT.
- [8] Wu, J. C. F y M. Hamada, Experiments, Wiley, New York, 2000.
- [9] Box, G. E. P., Hunter, W. G. y J. S. Hunter, Statistics for Experimenters, Wiley, New York, 1978.