

## **Metodología estadística para el estudio y evaluación de aspectos técnicos en la construcción**

**Pág. 21 - 33**

**Hernán de Solminihac T.**

Profesor del Departamento de Ingeniería de Construcción, Pontificia Universidad Católica de Chile, Casilla 6177, Santiago, Chile.

**RESUMEN:** este trabajo presenta una importante metodología estadística, denominada diseño de experimentos, que permite determinar al costo más bajo posible y con el mayor grado de confiabilidad, la efectividad real de nuevas técnicas, productos o procesos. Se describe, en una primera parte, los pasos principales para diseñar un buen experimento, enfatizando las actividades requeridas previas a la toma de los datos. La segunda parte muestra campos de aplicación de esta metodología en la Ingeniería de Construcción. Finalmente, se presenta el estudio de un ejemplo aplicado a la Ingeniería de Construcción.

### **I. INTRODUCCION**

La estadística puede ser una herramienta importante en las actividades que se desarrollan en la Ingeniería de Construcción, especialmente cuando se requiere determinar, al más bajo costo posible, la efectividad real de nuevas técnicas, productos o procesos. Esto se consigue realizando un seguimiento planificado, o diseñando un experimento específico, que permita analizar el efecto de las diferencias entre el proceso patrón y el proceso de interés. Esta metodología estadística, que se conoce como diseño de experimentos, permite estudiar los fenómenos a un nivel experimental, para luego, si el diseño fue bien realizado, inferir a producción los resultados de la experiencia.

El diseño de experimentos tiene como objetivo principal determinar la influencia de varias variables independientes (factores), en la variación de la variable dependiente (variable de interés). Este método permite obtener información útil con gran ahorro de tiempo y de recursos, especialmente cuando se compara con la alternativa de realizar varios experimentos con un sólo factor cada uno. Más aún, en un diseño se pueden incluir, además del efecto de cada factor, el efecto de una combinación de ellos (intersección).

Este trabajo describe los pasos principales para diseñar un buen experimento, enfatizando las actividades requeridas antes de tomar ningún dato. Además, se presentan campos de aplicación de esta metodología en el área de la Ingeniería de Construcción y se ilustra su utilización con el análisis de un ejemplo específico.

### **II. METODOLOGIA DE ESTUDIO**

#### **2.1 Generalidades**

Para el estudio de un fenómeno, es aconsejable seguir alguna metodología prefijada de modo de lograr los objetivos eficientemente. Este estudio recomienda seguir la metodología que se expone a continuación, la cual fue desarrollada en base a una combinación de los métodos propuestos por Anderson-McLean (1) y Haas-Hudson (2):

Paso 1: Reconocer que el problema existe. Esta, etapa consiste en tomar conciencia de que para mejorar una actividad, es necesario estudiar nuevos productos, o nuevos métodos, los cuales deben experimentarse antes de utilizarlos en gran escala.

Paso 2: Formulación del problema. Consiste en desarrollar una proposición de investigación, incluyendo sus objetivos, costos, beneficios y limitantes.

Paso 3: Diseño del experimento. Se diseña un experimento de modo de planificar eficientemente la toma de datos y su posterior análisis, ya que la forma de tomar los datos fija el método con que se debe analizar la información.

Paso 4: Toma de datos. El resultado positivo de una investigación experimental depende fundamentalmente de la validez de los datos obtenidos; por lo tanto, se debe tener un cuidado especial en este aspecto, tomando precauciones especiales como por ejemplo, diseñando un mecanismo de control y validación de la información.

Paso 5: Análisis de datos. El método de análisis a utilizar depende del diseño de experimento que se haya considerado.

Paso 6: Interpretación de los resultados. Esta etapa es muy importante en el proceso, ya que asocia el análisis estadístico a aplicaciones prácticas.

Paso 7: Elaboración de las conclusiones y recomendaciones. En esta última etapa se obtienen las conclusiones y se plantean las recomendaciones que se hayan desprendido del estudio.

A continuación se analizan, con mayor detalle, los pasos 3, 4 y 5; por considerarse que presentan una mayor complejidad desde un punto de vista estadístico.

## 2.2 Concepto general de diseño de experimentos

El objetivo principal de un diseño de experimento es determinar el efecto de varios factores (variables independientes), en algunas características de la variable de interés (variable dependiente).

Como se indica en la referencia (1), los aspectos más importantes a considerar para realizar un buen diseño de experimentos son: definir la variable dependiente que se va a medir; definir los factores que se van a usar en el experimento, con sus respectivos niveles; definir el espacio de inferencia para el problema; definir la cantidad de réplicas que se utilizarán para determinar el error; definir la metodología de análisis de datos y definir la metodología de toma de datos.

- a) Definición de la variable dependiente. Esta etapa, que al parecer es muy simple, permite definir claramente cuál va a ser la variable que se va a analizar. Hay veces que la definición de esta variable es más compleja, especialmente cuando se pretende estudiar una variable que es función de otras, tales como: promedios, desviaciones estándares, diferencias, etc. Por ejemplo, en el caso que se quiera comparar el proceso patrón con un proceso alternativo, una manera de estudiar este fenómeno es analizar el comportamiento de la diferencia entre ambas respuestas.
- b) Definición de los factores y sus niveles. El primer paso en esta etapa es determinar los factores o variables independientes que se estudiarán, es decir, el origen de la variación de la variable dependiente que no se puede atribuir al error experimental. De éstos, hay que elegir un número razonable de factores que se estime afectan dicha variación y que sean de fácil identificación en la práctica, ya que una cantidad muy grande de factores implicaría una gran cantidad de datos y por ende costos y plazos prohibitivos para el estudio.

Paralelamente con definir los factores, es necesario pensar en el número de niveles para cada factor. Los niveles definen el rango de variación de las variables independientes. Al igual que en el caso de los factores, el número de niveles debe estudiarse cuidadosamente, ya que muchos niveles implicaría la toma excesiva de datos.

Para determinar el número de factores y niveles en un experimento, existen varios criterios. Los dos principales son: incluir muchos factores con pocos niveles cada uno o incluir pocos factores con muchos niveles cada uno. El primer criterio es muy bueno cuando se trata de una investigación más general, en la cual se está tratando de eliminar estadísticamente algunos factores que en un principio parecen importantes. El segundo criterio, por el contrario, es útil cuando se trata de estudiar, una vez conocido su comportamiento general, el comportamiento específico (más detallado) de algunos factores.

El resultado final de esta etapa es definir el "factorial", que es una representación gráfica jerarquizada de la investigación. Un factorial está compuesto de factores (con sus respectivos niveles) y "celdas", que son un sector del factorial que abarca sólo a un nivel de cada factor. Una aplicación de estos conceptos se describe en el cuarto capítulo de este trabajo (ver Fig. N°1).

- c) Definición del espacio de inferencia. Este punto es muy importante, ya que se define el campo de aplicación de los resultados de la investigación. Es decir, se define si los resultados son aplicables sólo en un laboratorio, en una obra, en varias obras, en una ciudad, en un país o en todo el mundo. Si se desea dar una mayor universalidad a los resultados, se pueden incluir factores adicionales en el experimento.
- d) Definición de la cantidad de réplicas. Las réplicas son datos de características similares que se toman más de una vez, de modo de poder estimar el error involucrado en el experimento. En general, para el análisis de varianza la mínima cantidad de réplicas para determinar el error es uno, es decir, se deben tener dos datos que representen la misma información. Por otro lado, debido al costo involucrado en el ensayo normalmente no se consideran más de dos réplicas en los experimentos, es decir, tres datos equivalentes, ya que su aporte marginal a la explicación del fenómeno no es muy significativo para el análisis de varianza.
- e) Definición de la metodología de análisis de datos. El análisis de la información es posterior a la toma de los datos; sin embargo, debe estar completamente definido previo a la adquisición de la información, ya que la forma de toma de datos condiciona su análisis posterior. Si la información no puede adquirirse en la forma prevista, se debe ajustar el análisis a lo efectivamente realizado. El resultado final de esta etapa es la definición de un modelo matemático para el análisis de varianza.
- f) Definición de la metodología de toma de datos. Una vez definido el procedimiento de análisis, la metodología de toma de datos queda completamente determinada, ya que el análisis debe considerar factores tales como aprendizaje y experiencia, que de no ser considerados, las conclusiones obtenidas estarían sesgadas. Por ello, se debe definir claramente la secuencia en la cual se realizarán los ensayos. Esta importante relación entre la toma de datos y su forma de análisis, permite diferenciar los distintos tipos de diseños de experimentos. Si la secuencia es aleatoria, se tiene un diseño de experimento aleatorio y no existe problema de aprendizaje, ya que su efecto se distribuye al azar entre sus factores. Si la secuencia es por bloques, se tiene un diseño de experimento por bloques, el cual considera adecuadamente el aprendizaje. Además de estos dos diseños, existen otros de aplicaciones más específicas, que han sido muy bien descritos por Anderson y McClean (1).

### 2.3 Toma de datos

Una vez definida la metodología a seguir en el experimento, se debe iniciar la adquisición de datos. Esta etapa debe realizarse cuidadosamente, ya que una información mal tomada perjudicaría enormemente el resultado final de la investigación. En esta etapa, se deben definir sistemas de control que permitan validar la información. Algunas metodologías estadísticas para la adquisición de datos han sido descritas por Kohler (3) y por Clark y Shkade (4).

## 2.4 Análisis de datos

Los dos factores más importantes a considerar en el análisis de los datos son los siguientes:

- 1) Tipo de diseño de experimento. El diseño más fácil de analizar, pero más difícil de realizar en la práctica, es el diseño aleatorio, es decir, que la secuencia de toma de datos sea definida en base al azar. Otro tipo de diseño puede ser por bloques, en el cual la secuencia a realizar las experiencias es por bloques, y dentro de cada bloque la secuencia es aleatoria. Anderson y McClean (1) entregan una buena descripción de estos y otros tipos de diseño de experimentos.
- 2) Características del "factorial". Un factorial puede tener dos características: la primera es que sea balanceado, es decir, que el número de observaciones por cada celda del factorial sea igual. La segunda es que no sea balanceado, es decir, que el número de observaciones por cada celda del factorial sea diferente. Esta característica influye en forma importante en la operatoria matemática del análisis y, por lo tanto, debe considerarse en el estudio.

Una vez ejecutadas las experiencias, se deben analizar los resultados obtenidos. Para ello, una forma de hacerlo sería la recomendada en las referencias (5), (6) y (7):

- a) Crear una base de datos. Esta base debe crearse con toda la información disponible, definiendo muy bien los campos e introduciendo todos los registros adecuadamente.
- b) Verificar que las hipótesis del análisis estadístico se cumplen. Las hipótesis principales son: normalidad y homogeneidad de las varianzas. Existen diferentes métodos para realizar estas verificaciones. El método de Burr-Foster, conocido como el "Q-test" (1), entrega una metodología simple para determinar si las varianzas de las celdas del factorial son homogéneas. El método desarrollado por Shapiro y Wilk, conocido como "W-test" (1), presenta una metodología muy útil para determinar si un conjunto de datos tiene una distribución normal. Si los datos no cumplen con estas hipótesis, se debe realizar una transformación de los datos hasta satisfacerlas. Una precaución importante que se debe tener en este caso es con la interpretación de los resultados.
- c) Realizar un análisis de varianza (ANOVA) con el modelo previamente definido. El análisis de varianza es un filtro que permite determinar cuales factores tienen importancia estadística en la variación de la variable de interés.
- d) Análisis posterior. Una vez conocidos todos los factores significativos, si se desea, se continúa con el análisis; dependiendo de la investigación. Algunos ejemplos de esta etapa son: promedios, representaciones gráficas, regresiones, análisis de covarianzas, etc.
- e) Interpretación de resultados. Esta etapa es una de las más importantes, ya que aquí es donde se obtienen las conclusiones que deben tener una aplicación práctica.

## 2.5 Formas de disminuir la cantidad de datos

Existen algunos mecanismos que permiten disminuir la cantidad de datos a recolectar, sin perjudicar demasiado la contabilidad de los resultados y conclusiones obtenidas. Entre ellos se pueden mencionar los siguientes:

- a) Factores y niveles. Se puede pensar en disminuir el número de factores y niveles, manteniendo constantes para el experimento aquellos factores que no se van a considerar inicialmente en el estudio. Este procedimiento es bastante utilizado; pero, en general, disminuye el espacio de inferencia.

- b) Cantidad de réplicas. Existen dos maneras en que se puede pensar este punto: disminuir al mínimo la cantidad de réplicas en cada celda o sólo considerar réplicas parciales en algunas celdas del factorial. En este último caso estamos frente a un factorial no balanceado.
- c) Diseños de experimentos fraccionales. Estos diseños son experimentos que permiten disminuir en una cierta fracción el número de datos, siempre que todos los factores tengan el mismo número de niveles. El caso más común es considerar aquellos diseños que tienen factores con sólo dos niveles cada uno. Anderson y McClean (1 y 6) presentan un análisis detallado de los diseños de experimentos fraccionales.

### III. CAMPOS DE APLICACION

Esta metodología es aplicable en todas las áreas en las que se pueda tomar datos cuantitativos. En la construcción, tiene aplicaciones importantes en el control de calidad, la inspección de proyectos e investigaciones específicas. Dentro de estas últimas tenemos, por ejemplo:

- a) Las industrias de productos para la construcción. Estas industrias tienen en el diseño de experimentos, una herramienta muy poderosa para el estudio de sus productos. Ejemplos: aditivos, prefabricados, premezclados, etc. Dentro de este campo, las aplicaciones podrían ser: estudios de cambios de sistemas de producción, análisis de cambios en las materias primas, estudios de nuevos productos, etc.
- b) Comportamiento de materiales en obras de ingeniería, por ejemplo: hormigón, asfalto, etc. Algunos ejemplos en este aspecto podrían ser: estudios de cambios en las dosificaciones, del tiempo de curado, durabilidad o comportamiento a largo plazo, etc.
- c) Estudio del comportamiento de los trabajadores. Por ejemplo: estudio de rendimientos, comportamiento de la asistencia del personal a través del tiempo, etc.
- d) Estudio de los costos. Por ejemplo: análisis de variaciones del costo real versus el presupuestado, estudio de los componentes del precio unitario de una partida específica, etc.
- e) Comportamiento de los equipos. Por ejemplo, rendimientos de equipos, períodos de mantención, etc.
- f) Estudio de técnicas de construcción. Tales como: colocación de hormigón, construcción de moldajes, perforación de túneles, etc.

De estas actividades, donde probablemente es económicamente más factible aplicar la metodología propuesta es en la industria de productos para la construcción y el estudio del comportamiento de los materiales, ya que es posible intentar simular su comportamiento real en laboratorio. Los otros aspectos deben estudiarse en experimentos a nivel real, por lo tanto, el costo involucrado en el análisis es relativamente superior, pero se pueden, si se hace en diseño adecuado, obtener dividendos positivos a mediano plazo.

#### IV. ANALISIS DE UNA APLICACION ESPECIFICA

##### 4.1 Introducción

Para entender mejor la metodología del diseño de experimentos, este trabajo presenta el desarrollo de un estudio específico, que consiste en analizar los distintos factores que influyen en el resultado final o eficiencia (E) de una membrana de curado para hormigón.

Una forma de realizar este estudio sería ejecutar muchos ensayos con distintas probetas, sin tener un control estricto de cada factor. De esta forma se tendrían muchos datos para analizar el fenómeno, pero la inferencia que se puede obtener es poco confiable, ya que los datos podrían estar sesgados hacia algún factor, descuidando otros de igual o mayor importancia.

Otra forma de enfocar este problema sería en base a un diseño de experimento, que hecho en forma eficiente, podría incluso requerir menos datos y permitiría obtener conclusiones más confiables y aplicables hacia un espacio de inferencia más amplio.

Como vimos anteriormente, existen dos tipos de metodologías para realizar este estudio: analizar varios factores con pocos niveles cada uno o analizar pocos factores con varios niveles cada uno. A continuación se presentan ambos casos para el estudio de nuestro ejemplo.

##### 4.2 Estudio del caso de varios factores y pocos niveles

Si nuestro espacio de inferencia debe incluir las diferentes membranas existentes en el mercado, es natural que debemos incluir el factor "tipo de membrana" (M). Como vimos, el número de niveles debe ser mínimo en una primera etapa, digamos 2 niveles. Para incluir todo el rango comercial de calidades de membranas, tomemos los extremos de este rango, dejando para una segunda etapa, si se detecta alguna diferencia entre estos extremos, el análisis de los niveles intermedios. Si no detectamos ninguna diferencia entre el comportamiento de los extremos, lo más probable es que no se detecten diferencias entre las calidades de membranas localizadas al interior de este rango.

Probablemente, otro factor a considerar sería "tipo de hormigón" (H); de este modo se expandería nuestro espacio de inferencia, ya que de lo contrario nuestras conclusiones serían aplicables sólo a hormigones del tipo utilizado en la experiencia.

Otros factores que se pueden incluir son: "forma de aplicación" (F), "tiempo de aplicación" (T), "cantidad aplicada" (C), "temperatura ambiente" (A), "velocidad del viento" (V), "radiación solar" (R) y "humedad" (H).

Si una vez analizados los datos obtenidos en el programa experimental realizado, sólo se determina que el tipo de membrana (M), tiempo de aplicación (T) y cantidad aplicada (C) son los factores más importantes en la explicación de la variación de la eficiencia de las membranas de curado, y además, si deseamos conocer con más detalle el efecto de estos factores, se debe proceder a realizar la segunda etapa de este análisis, el caso de pocos factores y varios niveles.

##### 4.3 Estudio del caso de pocos factores y varios niveles

En este caso debemos realizar nuevamente el experimento, pero ahora considerando solamente los factores seleccionados en la etapa anterior y aumentando el número de niveles en cada factor. Si consideramos los tres factores seleccionados anteriormente con tres niveles cada uno, se podría construir el "factorial" mostrado en la Figura N°1, el que permite visualizar en mejor forma la magnitud del experimento.

##### 4.4 Metodología de análisis

Para el caso de los tres factores con tres niveles cada uno, suponiendo un diseño aleatorio y un factorial balanceado con una réplica en cada celda, el modelo que se indica en la ecuación 1 permite realizar un completo análisis de variancia para nuestro ejemplo:

$$E_{ijkn} = \mu + M_i + T_j + M^*T_{ij} + C_k + M^*C_{ik} + T^*C_{jk} + M^*T^*C_{ijk} + \varepsilon_{(ijk)n} \quad (1)$$

con: i: I, II y III ; j: 1,2, y 3; k: X, W y Z ; n: 1 y 2

donde:

$E_{ijkn}$	= eficiencia de la membrana de curado tipo "i", para el tiempo de aplicación "j", la cantidad aplicada "k" y el número de repeticiones "n"
$\mu$	= promedio general
$M_i$	= efecto del tipo de membrana de curado "i"
$T_j$	= efecto del tiempo de aplicación "j"
$M^*T_{ij}$	= efecto de la interacción de la membrana de curado "i" y el tiempo de aplicación "j"
$C_k$	= efecto de la cantidad aplicada "k"
$M^*C_{ik}$	= efecto de la interacción de la membrana de curado "i" y la cantidad aplicada "k"
$T^*C_{jk}$	= efecto de la interacción del tiempo de aplicación "j" y la cantidad aplicada "k"
$M^*T^*C_{ijk}$	= efecto de la interacción del tipo de membrana de curado "i", el tiempo de aplicación "j" y la cantidad aplicada "k"
$\varepsilon_{(ijk)n}$	= error aleatorio de la probeta número "n" para el tipo de membrana de curado "i", el tiempo de aplicación "j" y la cantidad aplicada "k". Donde "ε" tiene una distribución normal e independiente, con promedio cero y varianza igual a "σ <sup>2</sup> ".

## V COMENTARIOS FINALES

El diseño de experimentos es una herramienta muy útil para la evaluación de fenómenos técnicos en la construcción, especialmente en la industria de productos para la construcción y actividades de la construcción que son más repetitivas o de gran volumen.

Los pasos principales de esta metodología son: a) definición de la variable a estudiar (variable dependiente), b) definición de los factores y sus niveles (variables independientes), c) definición del espacio de inferencia, d) definición de la cantidad de réplicas, e) definición de la metodología de análisis de datos, f) definición de la metodología de toma de datos.

De estas etapas, a las que se le debe poner una atención especial son: conocer de antemano el espacio de inferencia a que se quiere llegar, evaluar con precaución la relación toma y análisis de datos y establecer claramente la cantidad de replicas que permitan determinar el error experimental.

Esta metodología considera partir de lo general para terminar en lo particular, es decir, investigar primero la influencia de muchos factores en poco detalle cada uno, para en una segunda etapa, una vez conocidos aquellos factores significativos en la variación de la variable dependiente, estudiar en profundidad el efecto de esos factores.

Además, dado que con el concepto de factorial se deben considerar normalmente muchos datos, existen mecanismos que permiten disminuir la cantidad de datos a tomar, sin perjudicar demasiado la confiabilidad y exactitud de las conclusiones obtenidas.

## REFERENCIAS

1. Anderson, V. L. y McClean, R. A., Design of Experiments: A Realistic Approach, Marcel Dekker Inc., New York. EE.UU., 1974.
2. Haas, R. y Hudson, W. R., Pavement Management System, R E. Kreiger Publishing Company, Florida, EE.UU., 1982.
3. Kohler, H., Statistics for Business and Economics, Scott, Foresman and Company, Illinois, EE.UU., 1985
4. Clark C.T. y Shkade L.L., Statistical Analysis for Administrative Decisions, South-Western Publishing Co., EE.UU., 1983

5. de Solminihac, H E., "Serviciability Rating of the Texas Highway System for Pavement Management". Master of Science Thesis, Department of Civil Engineering, The University of Texas at Austin, EE.UU., 1986.
6. de Solminihac, H E , Hudson W R. y Ricci, E., "Serviciability Rating of the Texas Highway System for Pavement Management", Research Report. 400-1F, Center for Transportation Research, The University of Texas at Austin, EE.UU., 1986.
7. de Solminihac, H.E., Hudson, W.R., y Ricci, E., "Metodología para determinar los Indices de Serviciabilidad de una Red Vial", VI Congreso Panamericano de Ingeniería de Tránsito y Transporte, Santiago, Chile, Diciembre 1986, pp. 17.
8. McClean, R. A. y Anderson, V. L., Applied Factorial and Fractional Designs, Marcel Dekker Inc., New York. EE.UU., 1984.

Tipo de Membrana (M)		I			II			III		
Tiempo de Aplicación (T)		1	2	3	1	2	3	1	2	3
Cantidad Aplicada (0)	X									
	W					 a <sup>n°j</sup>				
	Z									

Figura N° 1. Factorial utilizado para el estudio de la eficiencia de las membranas de curado.