



COMPUTACIONAL ESTADÍSTICA

INTRODUCCIÓN

Según Salmerón A. (2001), puede definirse la Estadística Computacional como la Estadística calculada. De esta manera, cualquier procedimiento estadístico que implique una gran cantidad de cálculos, puede considerarse como un procedimiento de Estadística Computacional.

Una de las ramas más destacadas dentro de la Estadística Computacional es la correspondiente a la simulación y métodos de Monte Carlo. Las técnicas estadísticas de simulación juegan un papel muy importante en las ciencias experimentales e ingenierías, como la Informática, Física, Química, Biología y, naturalmente, dentro de la propia Estadística.

La simulación es una técnica numérica para realizar experimentos en un computador, sobre cierto tipo de modelos matemáticos, de cara a hacer predicciones sobre el comportamiento de dichos modelos.

Podemos definir el método de Monte Carlo como una técnica para realizar experimentos (con la ayuda de un computador) sobre un determinado modelo para obtener información o realizar afirmaciones sobre los valores de ciertos parámetros de dicho modelo. En este contexto, entenderemos por modelo un conjunto de variables aleatorias y una serie de relaciones entre ellas.



Simulación

- Se trata de crear las condiciones que ocurren en un problema para ensayar una realidad.
- Es un estudio preliminar para luego crear un modelo matemático. Está basado en la modelización (estudio piloto).
- Es una técnica para ejecutar estudios pilotos, con resultados rápidos y a un costo relativamente bajo.
- La simulación de sistemas ofrece un método para analizar el comportamiento de un sistema.

Un sistema de simulación es una colección de entradas que pasan a través de las fases de cierto proceso, produciendo respuestas, por ejemplo: Sistema de manufacturas, sistema de servicios.

La simulación se describe comúnmente como un arte, o una ciencia sofisticada debido a que la utilidad de los resultados dependerá de la destreza del grupo que realiza y analiza el modelo. No existe una teoría científica para garantizar la validez de un proceso de simulación antes de que el experimento sea realizado. En su lugar la confiabilidad de un modelo es evaluada por la correspondencia de los resultados del modelo con los obtenidos por otros sistemas comparables con el que se está examinando.

Ventajas de la simulación:

- Da solución a problemas sin solución analítica.
- Generalmente es más barato mejorar el sistema vía simulación que hacerlo en el sistema real.
- La simulación proporciona un control total sobre el tiempo, debido a que un fenómeno se puede acelerar.
- Mediante los métodos de simulación es posible analizar sistemas de mayor complejidad o con menor detalle.



- Es mucho más sencillo utilizar y comprender los métodos de simulación que los métodos puramente analíticos.

Desventajas de la simulación:

- La simulación es imprecisa, y no se puede medir el grado de imprecisión.
- Los modelos de simulación en una computadora son costosos y requieren mucho tiempo para desarrollarse y validarse.
- Es difícil de aceptar los modelos de simulación y difícil de vender.
- Los modelos de simulación no dan soluciones óptimas.
- Requiere largos periodos de desarrollo.
- Surge el peligro de atribuir a los números un grado mayor de validez y precisión debido a que los resultados son numéricos.

Sistema: realidad

La situación real se lleva a un modelo, se pasa a desarrollar la metodología para estimar los parámetros y probar el modelo y luego se llega a la solución de la situación.

Estocástico: (probabilidad)

Todos los hechos están relacionados a elementos estocásticos y no es posible construir un modelo matemático, porque este modelo es determinístico.

Metodología de la simulación:

- Definición del sistema: equivale a definir la población sobre la que quiero realizar el estudio.
- Formulación del modelo ¿qué tipo de modelo voy a desarrollar?
- Colección de datos: revisar si hay datos históricos permitiendo obtener estimaciones preliminares.
- Implementación del modelo: llevarlo a la práctica.
- Verificación: Verificar que no haya errores durante la implementación del modelo.



- Validación: Se comprueba la exactitud del modelo, si está dando estimaciones reales.
- Documentación: Es el informe final (como se llevó a cabo la simulación).

No simular cuando:

- El problema se puede resolver usando “análisis de sentido común”.
- El problema se puede resolver analíticamente.
- No hay recursos disponibles para el proyecto.
- El modelo no puede ser verificado o validado.
- No hay información ni siquiera datos estimados.
- El comportamiento del sistema es demasiado complejo o no puede ser definido.

Modelación de un sistema:

El primer paso, es estudiar un sistema bajo un modelo el cual puede ser una explicación formal o una observación empírica.

Al analizar un sistema podemos observar que, al cambiar un aspecto del mismo, se producen cambios o alteraciones en otro, en estos casos la simulación es una buena alternativa.

Estructura de un modelo:

- Componentes
- Variables
- Parámetros
- Relaciones funcionales
- Restricciones
- Funciones de objetivos

Características deseables de un sistema de simulación:

- Que sea completo
- Adaptabilidad
- Credibilidad



- Simplicidad (menor número de parámetros)
- Factible, tanto en información como en recursos económicos

Clasificación de modelos:

- Modelos estocásticos: existe incertidumbre
- Modelos determinísticos: Modelos exactos

Otra clasificación:

- Modelos estáticos
- Modelos dinámicos

Comparación entre estadística tradicional y estadística computacional [Wegman, 1988]:

Estadística Tradicional	Estadística Computacional
Tamaño de muestra de pequeño a moderado	Tamaño de muestra muy grande
Conjuntos de datos distribuidos idéntica e independientemente	Conjuntos de datos no homogéneos
De una o baja dimensión	Alta dimensión
De cálculo manual	De cómputo intensivo
Preguntas bien enfocadas	Preguntas imprecisas
Fuertes suposiciones no verificables:	Supuestos débiles o no:
Relaciones (linealidad, aditividad)	Relaciones (no linealidad)
Estructura de los errores (la normalidad)	Estructura de los errores (de libre distribución)
Inferencia estadística	Inferencia estructural
Predominantemente algoritmos de forma cerrada	Posiblemente algoritmos iterativos
Optimalidad estadística	Robustez estadística