

# 전산통계

Chapter 1.3 모형

# 모형

---

- 자연 시스템은 수많은 구성요소가 있어 분석이 어려움
- 일반적으로 변수의 수가 매우 많아 연산횟수가 매우 커지고 시스템의 논리적 구성도 쉽지 않아 문제해결이 불가능
- **모형화 (modeling)**를 통해 시스템의 단순화 과정 필요
- **모형 (model)**은 단순화 과정을 통하여 자연 시스템의 구조를 나타낸 대응
- 모형이 하나의 식으로 표현될 때 이를 모형식이라고 함
- 일반적으로 모형은 여러 개의 식이 논리적 구조로 연결된 알고리즘 형태
- **따라서 효율적인 알고리즘을 구하는 것이 전산통계의 목적**

---

- 좋은 모형의 조건

- 실제 시스템의 주요 구성요소를 포함해야 한다.
- 환경에 맞추어 시스템 분석이 현실적으로 가능하도록 단순화가 이루어져야 한다.
- 모형에 의한 결과값과 실제 시스템의 결과값의 차이가 작아야 한다.

- 
- 모형화를 위한 기본적인 단계
    - 1단계 : 중요 구성요소의 선택
    - 2단계 : 모형의 종류와 형태 결정
    - 3단계 : 자료에 의한 모수의 추정
    - 4단계 : 모형에 의한 추정값 산출
    - 5단계 : 모형의 진단
    - 6단계 : 모형의 활용

---

- 분석자료의 조건

- 조사된 자료들이 현실세계에 나타날 수 있는 자료 전체를 대표하는 표본이어야 한다.
- 자료들 전체가 하나의 특성을 유지하고 있어야 한다.
- 미래를 예측할 때, 과거자료들이 보이고 있는 변화패턴이 미래에도 유지가 된다는 확신이 있어야 한다.

---

- 모형의 종류

- 형태에 따른 분류 (예 : 은행업무, 병원업무, 댐 건설, 매립)
  - 물리적 모형 : 실제와 동일한 모습으로 축소시킨 모형
  - 상징적 모형 : 논리적 표현이나 수식을 사용하여 구성된 모형
- 시간에 따른 분류 (예: 곤충연구)
  - 정적 모형 : 시간의 변화에 대하여 영향을 받지 않는 모형
  - 동적 모형 : 시간의 변화에 대하여 영향을 받는 모형
- 분포에 따른 분류 (예: 자료전송)
  - 결정적 모형 : 확률변수가 포함되지 않는 모형
  - 통계적 모형 : 확률변수를 포함한 모형

## 모의실험

---

- 모의실험 (simulation) : 실제 시스템을 모형화 한 가상적 모형을 수행하여 시스템의 변화 과정과 결과를 분석하는 실험.
- 결과에 이르는 과정을 볼 수 있는 장점
- 해석적 방법으로 최적해를 찾기가 어려울 때 사용되는 방법으로 인식
- 시스템의 구조와 종류에 대한 많은 지식과 논리화 과정 및 실험에 소요되는 시간이 요구되므로 이에 대한 세밀한 검토가 필요

---

- 장점

- 실제 시스템에서 자료를 얻기가 불가능하거나 비용이 많이 드는 경우나 파괴적인 실험이 불가피한 경우 모의실험이 선택
- 많은 변수들로 이루어진 복잡한 구조의 시스템 분석을 위해 사용하며, 특히 시스템의 구성요소들 사이의 내적 관계의 연구에 유용
- 외부환경의 변화와 내부 구성의 변화에 따른 시스템의 반응 및 효과를 분석할 수 있으며, 효과에 따른 구성요소들의 중요도를 알아볼 수 있으며 새로운 요소의 추가에 따른 문제점을 파악
- 복잡한 시스템의 변화를 파악할 수 있으며 최적 시스템의 구성조건을 알 수 있음



---

- 장점

- 확률과정의 변화현상과 같이 시스템의 운영을 동적 모형으로 사용 가능
- 모의실험은 구조적으로 파악하여 주요 구성요소를 선정하므로 실제 시스템보다 요소들의 제어에 손쉬우며 정밀도가 높은 결과를 도출 가능
- 새로운 사업의 계획단계 또는 공장자동화를 위한 새로운 공정의 도입 등 기존 시스템에 관한 정보가 거의 없는 경우 의사결정에 큰 도움
- 전체 시스템을 몇 개의 부시스템으로 나누어 분석하기에 쉬워 각 부시스템의 운용 전문가에 의하여 시스템을 모형화 가능

---

- 단점

- 일반적으로 모의실험은 수치해석 방법을 사용하므로 정확한 값보다 근사치 또는 수열을 제공하므로 결과분석에 유의하여야 함
- 새로운 실험을 위하여 많은 분석과 프로그래밍 시간이 필요
- 통계적 분포가 모의실험에 포함되는 경우에는 실제의 분포가 어떠한 형태인지 알기 위하여 많은 시간 동안 자료 조사 및 검토가 선행되어야 함

---

- 모의실험의 종류

- 사람과 기계 사이에 이루어지는 모의실험
- 프로그램으로 수행하는 모의실험: 통계적 모의실험

## 모형의 진단

---

- 오차 (error): 모형에서 고려되지 못한 요소들의 영향을 모은 것
- 잔차 (residual): 추정된 모형식으로 계산되는 추정값과 실제 자료의 차이
- 모형의 안정성: 훈련자료, 검정자료
- 모형의 타당성: 모형이 실제 시스템을 얼마나 잘 표현하고 있는지를 조사
- 유의사항
  - 모형에서 입력 자료 또는 가정된 분포가 실제와 얼마나 적합되는지 살펴보기 위해 실제 자료를 조사하여 적합도 검정 (goodness-of-fit test)을 통하여 검사
  - 변수의 형태와 유효자리수 등을 검사하여 절단 또는 반올림으로 인한 오차의 누적이 모형에 끼는 효과를 검사