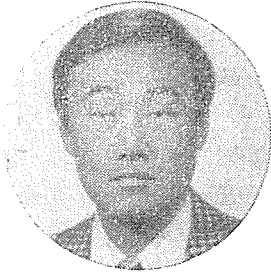


統計의 工業的 應用과 그 役割



서울대학교 전자통계학과 조교수

朴 聖 炫 박 사

1. 머 리 말

과학기술부문에 종사하는 사람들과 가끔 대화를 나누다 보면 이들중 통계학이 무엇이고 통계의 역할이 어떤가에 대하여 간혹 그릇된 지식을 가진 사람들을 보곤한다. 어떤이는 통계는 물가 통화량, 국민총생산과 같은 경제지수를 산출해 내는 학문으로만 생각하는 사람도 있고 어떤이는 통계는 확률을 다루는 수학의 일부분으로만 생각하는 분도 있다. 이들의 견해는 잘못된 것이며 통계학은 意思決定科學의 큰몫을 차지하는 학문으로 사회 및 자연의 온갖 현상을 연구하기 위하여 不確實性이 내포된 資料의 抽出, 分析 및 抽象을 통하여 의사결정에 필요한 정보의 획득 및 처리하는 방법을 연구하는 학문이라고 말할 수 있을 것이다.

자료처리 및 의사결정과학으로서 통계학의 기여는 인문, 사회, 경제, 자연, 의학 등 광범위한 영역에 걸쳐 공헌하고 있지만 이 論說에서 주로 다루고자하는 것은 통계가 공업 및 공업발전을 위해서 어떻게 기여하고 또 기여할 수 있는가에 대해서 얘기해 보고자 한다. 특히 현재 우리가 당면하고 있는 第4次 經濟開發 5個年計劃에 있어서 과학기술부문의 기본목표가 기술혁신과 능률의 향상, 산업의 질적인 고도화에 있다고 보겠는데 統計科學技術도 유용하게 적용하므로써 공산품의 질적향상, 새로운 공업기술의 개발, 기획 및 관리의 과학화등에 도움을 줄 수

있다면 제4차 경제개발계획의 달성을 위하여 뛰어나 보탬이 될 수 있는 일일 것이다.

통계가 공업부문에 어떻게 쓰여지고 있는가를 크게 두가지로 나누어서 설명하고자 한다. 첫째로 자료처리의 諸方法을 통한 통계의 공업적 응용과 둘째로는 실험계획법 및 분석을 통한 응용이다. 그리고 이 글을 맺기 전에 운영과학이라고 불리워지는 Operations Research(OR)와 전자계산학(Computer Science)과의 유대관계를 밝히고 공업발전을 위해서 상호 긴밀한 협조가 바람직함을 얘기하려 한다.

2. 資料分析의 諸 技法

공업생산품을 만들어내는 대부분의 제조공장에서는 적은 양이든 많은 양이든 간에 여러종류의 자료를 계속해서 획득하고 있다. 특히 생산과정중에서 품질관리, 안전관리 등에 관한 자료와 화학공장일 경우에는 화학반응조건등에 관한 자료들이다. 이런 자료들은 대부분 무시할 수 없는 誤差를 내포하고 있으며 또 눈에 잘 띄지 않는 여러 要因들의 함수관계를 가지고 있으므로 이들 자료가 줄 수 있는 최대의 情報를 얻기 위해서 올바른 자료처리방법을 선택하는데 세심한 주의가 필요하다. 실제로 비용을 많이 들여 자료를 얻어놓고 올바른 분석방법을 몰라 중요한 자료를 무용하게 방치시키는 예가 허다하다. 공업에서 관리기술의 질적 향상과 능률의 극대화를 위해서 올바른 자료처리방법을 이용할 수 있

는 능력은 매우 바람직하다. 다음에서 보통 많이 쓰이는 몇가지 기법을 간략하게 설명하고자한다

(1) 記述的인 方法

가장 기초적인 것으로 주어진 자료를 Histogram으로 표시하거나 평균치를 구하거나 표준오차를 계산하거나 하는 방법이다. 간단명료하게 자료를 분류할 수 있으면 간혹 새로운 사실을 발견할 때가 있다.

(2) 分散分析 및 假說檢定

誤差의 소재를 밝혀내고 변수들이 주는 영향을 알아내기 위해서 分散分析은 필요하며, 어떤 주어진 假定이 옳은 것인지 그른 것인지를 자료로부터 판정하기 위해서는 假說檢定方法이 쓰여져야 할 것이다. 새로운 공업기술을 개발할 때에 아직까지 써오던 기술보다 우수한가 아닌가 하는 것은 철저히 檢定해 볼 가치가 있다.

(3) 相關分析 및 回歸分析

여러 변수간의 상관관계 및 반응관계를 규명하기 위해서 상관 및 회귀분석은 매우 적절한 방법이다. 이 분석방법은 널리 사용되고있는 기법으로 여러 要因이 상호작용하여 어떤 함수관계를 가질때에 이를 효과적으로 측정하는 방법을 제시해 준다. 예를들어 어떤 화학반응량이 온도 압력 및 습도의 함수라고 하자. 만약 실험을 통해서 통계적으로 이 함수관계를 규명한다면 품질의 관리 및 향상을 위해서 큰 자료를 얻을 수 있을 것이다.

(4) 管 理 圖

제품을 생산하고있는 과정에서 제품의 품질을 계속적으로 검토하기 위하여 여러종류의 품질관리도를 사용하고 있는데 그중 Shewhart 관리도나 Cumulative sum 관리도 같은 것이 흔히 쓰인다. 이 도표들은 제품의 생산과정이 통계적으로 안정상태에 있는가 없는가를 보여줄 수 있기 때문에 매우 유용하게 이용된다.

(5) 標本抽出法

표본을 사용하여 자료를 추출할 때에 어떻게 표본을 선택할 것인가에 대하여 세심한 주의를 하지않으면 이 표본으로 이루어진 분석이 精度가 크지 못하고 가치가 없을 수 있다. 따라서유용한 통계적분석을 얻기 위하여 올바른 표본 추출법을 사용하여야 한다. 또한 표본추출법에 따라서 분석방법도 달라져야 할 경우가 많다.

(6) 통계적 試무레이슨

기술을 개발하거나 어떤 문제를 풀기 위해서 공장의 쉐시시스템을 상대로 실험하는 것이 시간과 경비관계로 힘들 경우가 많다. 이럴 때에 실제의 쉐시시스템을 흉내내서 수치적으로 試무레이슨 모델을 만들고 관련된 변수들의 통계적인 분포에 따라서 난수(random number)를 발생시켜 이 모델이 어떤 결과를 주는가를 관찰하므로써 실제 쉐시시스템의 결과를 가상적으로 분석하는 것이다. 최근에 많이 쓰여지는 분석방법으로 거의전 자체산기를 이용하여 하고 있다.

3. 實驗計劃法과 分析

앞에서는 주로 주어진 자료로부터 최대의 情報을 얻기 위한 통계적분석방법을 열거하였는데 실제에 있어서 이들 자료를 획득하기 전에 자료 획득을 위한 방안이 체계적으로 되어있지 않으면 간혹 얻어진 자료가 무용지물일 경우가 있다 경제적 면을 고려하여 적은 양의 자료로부터 많은 정보를 얻기 위해서는 자료획득을 위한 올바른 실험계획법이 수립되고 이에 따라서 자료의 획득 및 분석이 이루어져야 한다. 분석방법도 실험계획에 따라 달라지기 때문에 처음부터 어떠한 목적을 위하여 자료를 구할려는가를 뚜렷이 설정하고 분석방법을 구상한 후에 이들을 만족시킬 수 있는 실험계획법을 정하여야 할 것이다.

실험계획법을 두가지 형태로 분류할 수있는데 한가지는 실험단위의 배치에 따른 분류이고 다른 하나는 處理組合의 선택에 따른 분류이다.

배치에 따른 분류로는 흔히 쓰이는 것이 完全任意配列法(Completely random design), 亂塊法(Randomized Complete block design), 라틴方格法(Latin Square design), 不完備블럭法(Incomplete block design) 등을 들을 수 있으며 처리조합에 의한 분류로서는 要因實驗(Factorial experiments)이나 部分要因實驗(Fractional factorial experiments)을 주로 들을 수 있다.

예를 들어 어떤 품목을 생산하는데 두개의 상이한 생산과정이 개발되었다고 하자. 이를 실험을 통해서 검토할 때에 만약 3反復 실험을 한다면 이것을 要因實驗의 입장에서는 單一要因(水準이 2인 單一因子)에 관한 실험이 될 것이다. 그러나 이 실험을 3개의 블럭을 이용한 亂塊法으로 실험단위에 중점을 두어서 처리할 수도 있을 것이다. 물론 어떤 실험계획법을 택할 것인가 하는 것은 소기의 실험목적에 달성하려는 분석 방법등을 고려한 후에 정해져야 할 것이다.

마지막으로 첨가시키고 싶은것은 최근에 화학공장들을 중심으로 탐지하려는 실험계획법이 소개되고 있는데 예를들면 中心合成計劃(Central composite design)이나 回轉計劃(Rotatable design)같은 것이 있다. 이들 실험계획법을 反應表面分析을 위한 計劃法이라고도 흔히 불리워진다.

4. 맺 는 말

아직까지 공업부문의 제조공장을 중심으로 자료처리를 위한 통계의 여러 분석방법과 자료의 효율적인 획득을 위한 실험계획법에 관하여 간략하게 설명하였다. 자료의 획득과 분석을 통한

(13p에서 계속됨)

선진국과 기초과학의 수준의 차가 어느정도 단축은 되기는커녕 나날이 커지고 있을뿐더러 국내의 공업 경제 발전과 비해서도 기형적으로 그 칙체성을 지니고 있음을 말하였다. 기초과학 연구의 발전은 고도의 공업수준을 유지한다는 이유에서 뿐만아니라 우리가 불의에 당면할는지도 모르는 환경문제 에너지 문제등의 해결에 도움이 될것이고 참다운 자연과학교육의 든든한 기

통계의 공업적 응용과 그의 역할은 실로 매우 크다고 하겠다. 품질의 엄격한 통계적관리를 통한 질적향상 및 불량품의 감소를 꾀할 수 있고 생산과정중에 관련된 여러 변수간에 함수관계를 통계적으로 추정하여 반응의 예측 및 공정기술의 개량이 가능할 것이다. 기타 수없이 많은 응용 범위가 있으며 일단 자료를 취급처리하는 업무는 통계업무로 간주되어야 할 것이다.

마지막으로 최근에 의사결정과학의 분야로서 널리 보급되기 시작하는 OR과 계산학과 통계학의 유대관계를 밝히고 어떻게 상호협조되는가를 말하고자 한다. OR의 諸 技法中에는 확률적 모델을 많이 사용하고 있는데 在庫理論, 待期理論, 信賴度理論, 確率過程論등을 들을 수 있으며 이들은 통계의 중요한 개념인 확률을 바탕으로 한 이론들이므로 OR와 통계학과의 관계는 입술과 이와의 관계같다고 할 것이다.

과학과 기술이 급성장함에 따라서 자연히 대량정보의 처리와 분석이 불가피하여졌는데 이는 대부분의 통계적자료분석이 전자계산기의 도움이 필수불가결하게 된 것을 의미한다. 이처럼 통계업무처리와 전자계산기를 효율적으로 사용하는 기술은 배어놓고 생각할 수 없게 되었다. 현재 서울대학교에는 계산학과 통계학이 한 課로 묶여있는 사실만 보아도 알 수 있을 것이다.

바람직스러운 일은 각종의 공업기업체에서 품질의 관리 및 향상, 기술개발등을 원활히 하기 위하여 필요한 자료의 획득계획, 분석등을 효율적으로 할 수 있는 統計人을 가지고 있어야하며 성격상으로 통계업무가 OR이나 電算계통과 밀접한 관계가 있는 만큼 같은 課로 배속시켜 企劃部나 研究開發部에 두는것도 좋은 생각일 것이다.

반을 마련한다는 면에서도 모든 기초분야에 균형있고 알맞는 규모의 인적자원과 시설을 갖출 것을 목표로 삼아야한다. 끝으로 과거 10년 동안의 심한 응용과학 위주의 방침은 하루속히 수정되어야 하며, 국가의 장기적이고 일관성 있는 기초과학육성의 새로운 계획수립이 필요함을 말하였다.

Naval Research Laboratory
Washington, D.C. 20375