*Proceedings of Joint Conference of Korean Data And Information Science Society and The Korean Data Analysis Society, April 30 - May 1, 2004, pp. 167-174*

# 퍼지이론의 통계적 응용을 위한 교과목 연구*

이영섭 · 김혜중[1] · 이주성 · 안성헌[2]

## 요 약

Zadeh(1965)에 의하여 도입된 퍼지이론은 최근 컴퓨터공학이나 산업공학에 응용되기 시작하면서 그 유용성이 확인된 후 여러 분야에서 관심을 갖기 시작한 새로운 이론이다. 특히 제 산업분야에서 나타나는 통계모델의 정확한 분석을 위한 퍼지이론의 이용은 그들 분야의 발전은 물론 새로운 통계분석 방법을 제시하는데 큰 의의가 있다하겠다. 이와 같은 중요성에 비추어 퍼지이론을 이용한 통계분석을 학생들에게 효과적으로 학습시키는 것은 매우 중요한 일로서 이 연구는 통계분석방법을 퍼지이론으로 이해하고 또한 새로운 통계적 퍼지 모델을 어떻게 개발하고 응용할 것인가를 제시하고자 하는 교과목 연구이다. 이 연구가 향후 다양한 시대적 요구에 부응하는 새로운 교과목 개발의 전기가 되기를 기대한다.

## Ⅰ. 서론

Zadeh(1965)에 의하여 도입된 퍼지이론은 최근 컴퓨터공학이나 산업공학에 응용되기 시작하면서 그 응용성이 확인된 후 여러 분야에서 관심을 갖기 시작한 새로운 이론이다.

최근 방대한 자료를 컴퓨터를 이용하여 처리함에 있어 공간적인 자료나 다차원의 자료분석에 대한 중요성이 증가하고 있다. 이러한 자료들을 효과적으로 분석하기 위하여 측도에 관한 다양한 이론들과 분석기법들이 개발되어 활용되고 있다. 이러한 시도 중에서 1994년 Sugeno에 의해서 도입된 퍼지 측도 및 퍼지 적분론은 퍼지이론이 통계학은 물론 다른 응용과학에 접목시킬 수 있는 매우 유용한 분야 중의 하나임을 인식시켜 주었다(Zadeh, 1975)

특히 제 산업분야에서 나타나는 통계모델의 정확한 분석을 위한 퍼지이론의 통계학에서의 응용은 그들 분야의 발전은 물론 새로운 통계분석 방법을 제시하는데 큰 역할을 하고 있다. 즉, 퍼지 집합(fuzzy set)이나 퍼지 논리(fuzzy logic)에 기초한 다양한 통계적 모형이나 방법들이 학계나 산업분야( 자료분석이나 시스템 제어 분야 등) 에서 개발되고 그 이용이 점차 증가하고 있다. 최근 퍼지 자료의 통계적 추론, 퍼지 순서 통계량(fuzzy order statistics), 그리고 이러한 추론이나 통계량을 이용한 퍼지 군집분석(fuzzy clustering)등이 연구되고 있으며 이들은 이용하여 이상값을 찾거나 로버스트 추정 등을 시도하고 있다.(Manton et al., 1994)

---

*이 연구는 한국학술진흥재단 [대학교육과정 개발연구 지원사업]과 동국대학교의 지원으로 이루어졌음
[1] 100-715 서울시 중구 필동 3가26 동국대학교 통계학과
[2] 100-715 서울시 중구 필동 3가26 동국대학교 수학과

이와 같은 중요성에 비추어 퍼지이론을 이용한 통계분석을 학생들에게 효과적으로 학습시키는 것은 매우 중요한 일로서(안성헌 외, 2003) 이 연구는 통계분석방법을 퍼지이론으로 이해하고 또한 새로운 통계적 퍼지 모델을 어떻게 개발하고 응용할 것인가를 제시하고자 하는 교과목 연구이다. 특히 학생들에게 이와 같은 과목을 이수하게 함으로서 졸업 후에 현업에 빠르게 적응할 수 있는 능력을 배양시키고자 하는 것이 본 연구의 목적이다. 이 연구가 향후 다양한 시대적 요구에 부응하는 새로운 교과목 개발의 전기가 되기를 기대한다.

## Ⅱ. 본론

### 2.1 교과목의 개요

본 교과목은 퍼지이론에 대한 소개와 함께 통계학에서의 퍼지이론의 통계적 응용을 위한 다양한 통계모델을 제시하고 이를 학습함을 목표로 한다. 그러므로 이 강좌를 통하여 퍼지이론의 수학적 기초, 퍼지 집합의 이해, 퍼지와 crisp이론의 비교, 퍼지 측도 및 적분의 소개 및 응용을 심도있게 학습한 후 통계분석을 위한 퍼지사상의 확률모델, 퍼지자료를 위한 GOM 모델, 퍼지베이지 결정모델들을 소개하고 , 퍼지 순서 통계량에 기초한 로버스트 추정과 퍼지 군집분석등 다양한 통계적 응용기법들을 익히도록한다.

### 2.2 교과목의 목적

이 연구는 통계분석방법을 퍼지이론으로 이해하고 또한 새로운 통계적 퍼지 모델을 어떻게 개발하고 응용할 것인가를 제시하고자 하는 교과목 연구이다. 특히 학생들에게 이와 같은 과목을 이수하게 함으로서 졸업 후에 현업에 빠르게 적응할 수 있는 능력을 배양시키고자하자 하는 것이 본 연구의 목적이다. 이와 같은 목적을 달성하기위한 세부적 목적은 다음과 같다.

1. 퍼지이론과 퍼지 논리의 기본적인 개념을 파악하게하며,
2. 퍼지이론의 응용분야로써 퍼지 통계학의 이해와 통계모델 구축 기법등을 습득 하게하고,
3. 퍼지 통계학을 충분히 이해하여 제 분야로 확장시켜 응용할 수 있는 능력 습 득하게 한다.

*** 수강신청시 유의사항 (선수과목 제도)**

본 강좌는 수학과 통계학의 기초 과목들을 이수한 학생으로서 특히 집합과 논리이론, 확률 및 통계학, 선형대수, 프로그래밍과 알고리즘에 대한 지식이 있어야만 수강 할 수 있는 학부 4학년 또는 대학원 석사과목이다.

### 2.3 강의 계획서

본 강좌는 원래 목적이 수학과 통계학의 접목 교과목으로써 퍼지이론과 퍼지 통계학으로의 응

용에 관하여 학습하는 교과목이다. 따라서 강의는 크게 두 부분으로 나누어서 팀티칭 형식으로 16주간 이루어진다. 전반부에서는 퍼지이론의 기본개념들에 대하여 강의하고 후반부에서는 퍼지 사상의 확률모델 및 퍼지이론의 통계적 응용을 중심으로 강의한다.

각 주별 강의 내용은 다음과 같다.

**1주: 퍼지개념의 통계적해석의 이해**

퍼지이론의 역사를 소개하고 퍼지개념과 크리스프개념을 비교 공부한다. 또한 통계학에서의 확률변수와 퍼지이론에서 퍼지수의 개념의 차이를 설명하고 이들 사이에서 일어나는 실제문제들에 대하여 다양한 예제들을 소개함으로 학생들에게 퍼지이론의 필요성에 대한 이해를 돕도록 한다.

**2주: 퍼지집합(Fuzzy sts)의 표현과 퍼지집합의 연산**

퍼지집합과 멤버쉽함수를 정의하고 이들의 연산에 대하여 공부한다. 또한 퍼지집합의 다양한 연산들을 소개하고 그 성질 등을 공부한다. 특히 퍼지집합들의 대수적 합집합과 대수적 곱집합을 정의한다. 이들의 성질에 관한 증명은 멤버쉽 함수의 정확한 이해에서 비롯된 것임으로 이들의 심도있는 학습이 필요함을 주지시킨다.

**3주: 멤버쉽함수(Membership Function)와 멤버쉽함수의 구성**

퍼지개념이 주어 질 때 그 개념을 나타내는 멤버쉽함수를 정의하는 방법에 대하여 공부한다. 객관적 부분과 주관적 부분으로 이루어저 있는 퍼지개념중 특히 객관적부분의 중요성을 이해시키고 이를 구성하는 방법(평균법, 위치파라메타 추정법)들을 공부한다. 또한 멤버쉽함수의 분포와 통계학에서의 평균과 분산에 의한 분포의 차이를 설명한다.

**4주: 확신구간의 연산과 퍼지수의 연산**

확신구간의 4칙 연산과 퍼지수의 4칙 연산에 대하여 공부한다. 퍼지수의 4칙 연산은 $\alpha$-레벨 구간의 연산에 기초함으로 $\alpha$-레벨구간에 의하여 표현된 두 퍼지수에 대한 연산에 대하여 심도있게 학습한다. 또한 삼각형 퍼지수의 연산법을 공부하고 사다리꼴 퍼지수의 연산, L-R 퍼지수의 연산에 대하여 학습한다.

**5주: 일반관계와 퍼지관계의 이해**

보통관계의 확장으로서 정의된 퍼지관계에 대하여 공부한다. 이를 위하여 일반관계를 간략히 소개하고 퍼지관계를 엄밀히 정의하며 이들의 연산법과 퍼지관계를 행렬로 표시하는 방법을 소개한다. 특히 퍼지관계에서 유사관계와 이 관계에 의하여 유도되는 동치류에 대하여 다양한 통계적 모델들을 예로 제시한다.

**6주: 퍼지사상과 통계적 모델**

퍼지관계의 특수 관계로 퍼지사상을 정의하고 이 사상에 의한 퍼지집합의 상과 역상을 공부하며 본 강좌의 후반부에 다루어질 다양한 통계모델을 소개한다.

**7주: 일반추론과 퍼지추론**

일반 이론의 증명을 위한 일반추론법칙을 간단히 소개하고 퍼지추론방법(일반적으로 근사적 추론이라 함(Approximate reasoning))을 공부한다. 특히 일반추론과 퍼지추론방법의 차이점을 정확히 이해시킴으로서 학생들이 퍼지이론의 논리전개에 도움을 주도록 한다.

**8주: 퍼지측도와 퍼지적분**

불확실성의 수치적 표현을 위한 확률의 확장된 개념으로서 퍼지측도를 소개하고 다양한 통계모델을 제시한다. 또한 퍼지측도를 이용한 퍼지적분을 정의하고 퍼지적분을 평가에 응용한다. 퍼지적분의 응용에의 평가는 개개의 평가와 퍼지측도에서 종합평가를 구하는 일과 개개의 평가와 종합평가에서 퍼지측도를 구하는 문제로 대별된다. 특히 의사결정문제에서 퍼지적분에 의한 종합평가 방법에 대하여 공부한다.

**9주: 「퍼지 사상의 확률 모델」**

앞에서 배운 퍼지이론에 근거하여 확률적인 애매함을 포함하는 베이지론적인 의사결정문제에서 인간의 주관적인 애매함의 상태로서의 사상 그 자체에 포함되는 경우의 문제를 다룬다. 그러기 위하여 Zadeh에 의하여 도입된 퍼지사상의 확률에 대하여 설명한다. 궁극적으로 퍼지 사상계의 확률모델은 보통 사상계의 확률모델과 형식적으로 거의 일치하고 있음을 주지시킨다.

**10주: 「퍼지집합 모델 또는 GOM(Grade of Membership)모델의 모수추정」**

퍼지 집합에서 어떤 요소 $i$가 $k$번째 집합에 속하는 정도를 나타내는 GOM점수를 정의하고 GOM점수에 따라 통계학에서의 crisp와 퍼지 집합을 정의 한 후 통계적 모수추정에 대해서 알아본다. 한편, GOM모델은 퍼지 집합 모수를 사용하여 이산형 자료(discrete data)가 어떻게 실험결과의 확률 $p$ 와 연결되는지를 말해준다. 이를 위하여 퍼지 집합에서의 우도 방정식(likelihood equation) 방법에 대해서 알아보고 이 우도함수의 변형된 형태와 이산 요인 분석(discrete factor analysis)과 같은 이산형 반응 변수 자료에 대해 적용하는 다른 다변량 모형과의 비교에 대해서 알아본다. 마지막으로 Latent Class Models(LCM)과 GOM모델을 비교학습한다.

**11주: 「집단화 된 자료(aggregate data)에서의 GOM 모델」**

이제까지 개인 단위의 다변량 이산 반응 자료를 분석하는 기본적인 모델에 대해서 알아보았다.

본 강의에서는 개인자료가 합쳐진 집단화된 자료를 다루기 위하여 중요한 모델 변형에 대해서 공부한다. 먼저 포아송모델, 음이항모델에서 집단화된 자료의 우도함수와 모수 추정과 GOM점수와의 관계를 알아본다. 즉, GOM 우도함수의 포아송 변형을 사용하여 그룹이나 집단화된 사건비율을 분석함으로써 퍼지 집합을 도출하는 GOM모델을 소개한다.

### 12주: 「퍼지 베이지 결정 모델」

베이지 결정 모델에서 상태나 정보가 퍼지화된 경우의 결정모델에 대하여 알아본다. “대개 어떠한 상황에 처하게 되면, 어떤 결정을 취하면 되는가?”와 같은 애매한 상황에 있는 결정 문제를 다룬다.

### 13주: 「퍼지 측도의 응용」

불확실성의 측도로 많이 이용되고 있는 Shannon의 엔트로피에 대해서 공부한다.. 이것은 통계학의 많은 분야에서 불확실성이나 불순도의 측정도구로써 많이 응용되어 사용되어지고 있다. 또한 거리측정에 대한 측도를 학습하기 위하여 다양한 모델들을 제시한다.

### 14주: 「퍼지 신경망 모델(Fuzzy Neural Networks」

신경망 분석에 대해서 간단히 설명하고, 퍼지 신경망 모델의 구조에 대하여 도표와 함께 설명한다. 이 모델은 신경망 분석에서 퍼지시스템의 기능을 도입한 것으로써 crisp한 입력변수가 퍼지성(fuzzification)을 갖고서 다음 단계(layer)로 퍼지 법칙(fuzzy rule)에 의해 전달되고 마지막 출력단계에서 비퍼지화(defuzzification)과정을 거쳐서 crisp 한 출력 자료를 제공하는 모델이다.

### 15주: 「퍼지 신경망 모델의 응용」

Buckely와 Hayashi(1994)의 논문을 중심으로 퍼지 신경망 모델의 응용에 대해서 강의한다. 퍼지 자료 사이의 함수적인 관계를 알아보는 퍼지 회귀(fuzzy regression),주로 경제학에서 사용되는 퍼지 행렬식을 주는 과정, 그리고 마지막으로 퍼지 분류시스템(fuzzy classification system), fuzzy genetic algorithm에 대해서 학습한다.

### 16주: 「퍼지 순서통계량과 퍼지군집분석으로서의 응용」

통계학에서 중위수와 중위절대편차(median absolute deviation, MAD)는 순서 통계량(order statistics)에 기초한 로버스트 통계량이다. 이 순서통계량을 퍼지 집합으로 확장시켜 퍼지 중위수와 퍼지 MAD를 정의한다. 퍼지 $k$ 평균 군집분석 알고리즘(fuzzy $k$-means clustering algorithm)은 $k$ 평균 군집분석 알고리즘의 군집정의를 퍼지 개념을 도입하여 군집에 속할 가능성을 퍼지 확률로써 정의하여 군집분석을 하는 것을 말한다. 이에 대한 자세한 알고리즘과 응용에 대해서 학습한다.

## 2.5 과제물과 프로젝트

과제물은 매주 주어지며 특별한 언급이 없는 한 과제 부여일로부터 일주일 후에 수업시간 전에 제출하는 것을 원칙으로 한다. 전반부에서는 주로 퍼지 이론에 대한 전반적인 이해를 증진시키기 위하여 이론의 증명 중심으로 과제물을 부과한다. 후반부에서는 통계적인 기초 이론의 증명 및 응용뿐만 아니라 실제 자료를 이용하여 컴퓨터 프로그램을 이용한 실습도 과제물로 번갈아 가면서 주석과 함께 다시 돌려준다. 그 이외 자세한 정답은 과목 홈페이지에 게시하도록 한다. 프로젝트는 학기 중간에 주어지며 학기 말 까지 제출하도록 한다.

프로젝트는 실제 자료를 수강생들에게 주어서 통계적 응용 기법을 활용하여 분석하도록 한다. 특히 퍼지 $k$ 평균 군집분석을 적용하여 일반적인 $k$ 평균 분석기법과 결과를 비교 분석하도록 한다. 아주 우수한 프로젝트는 자동적으로 A 학점을 받도록 한다.

## 2.6 강의 자료 및 교재

전반부의 퍼지 이론에 대한 설명에서는 한글교재를 주로 사용하며 후반부의 통계적 응용에서는 적합한 한글 교재가 없기 때문에 주로 외국 원서나 논문들을 참고로 하여 유인물을 만들어 수업을 진행한다. 다음과 같은 교재를 주 교재로 사용하나 필요할 때 마다 다른 교재 일부나 논문 등을 학생들에게 부교재로 소개하여 강의 내용을 학생들에게 충분히 이해하도록 한다.

① 퍼지 이론(1주에서 8주) : 장이채(1998), 유동선,이교원(1998), 박민용, 최항식(1990).

② 퍼지이론의 통계응용(9주에서 16주) : Manton et al.(1994), Negoita and Ralescu (1987), Buckley and Hayashi(1994).

그 외의 모든 강의 자료는 과목 홈페이지에 게시하도록 한다.

## 2.7 평가방법

최종 학점은 출석, 과제물, 중간고사, 기말고사, 프로젝트, 등을 기준으로 평가하는 것을 원칙으로 한다. 모든 평가물에 대해서는 수강생들이 자신의 점수를 충분히 인식하도록 주석과 자세한 설명을 하여 의문사항이 없도록 한다.

평가요소들의 비율은 출석 10%, 과제물 20%, 중간고사 20%, 기말고사 30%, 프로젝트 20%로 한다. 이러한 비율을 기준으로 학기말에 100점을 기준으로 환산하여 상위 학점 (A,A+)의 비율을 최대 20%로 제한하여 수강생들끼리 상대 평가를 실시한다. 이러한 상대 평가로 인하여 보다 적극적으로 수업에 참여하고 경쟁력을 갖추게 한다.

## Ⅲ. 결론

퍼지이론과 통계학을 접목하여 새로운 교과목을 개설한다는 것은 수리과학의 발전을 위하여 아주 의미있는 일이라고 할 수 있다. 퍼지이론이 단지 이론에만 머물러 있지 않고 다른 학문과의 연계를 통하여 기존 이론을 보다 확장하고 실생활에 더 가까이 응용될 수 있기 때문이다. 퍼지 이론의 응용은 비단 통계학 뿐 아니라 인공지능, 의료진단시스템(이광형,오길록,1991), 고로제어(박용선, 최한식, 1990)등 다양한 분야에서도 응용되고 있다. 최근 이러한 분야들 중에서 확률개념을 퍼지집합이나 퍼지논리로 확장된 퍼지개념을 자료분석이나 시스템 제어 분야에 적용하여 많은 좋은 결과들을 얻어내고 있다.

퍼지 이론과 응용을 한 학기 3시간 강좌로 모두 강의하기에는 무리가 있을 것이라 생각된다. 이를 감안하여 본 연구에서 제안한 강의 내용과 계획은 기본 퍼지이론과 이의 통계적 응용에 국한하였다. 그러므로 수리과학을 전공하는 학생들이 본 강좌를 성실히 따른다면 수강생들이 졸업 후에 퍼지 이론을 현업에 사용하는 데는 큰 무리가 없을 것이라고 확신한다.

전공간의 상호 세미나나 워크샵을 통하여 서로의 학문을 이해하고 서로 응용할 수 있는 교과목은 무한히 많다. 특히 현대 산업사회에서는 학문간의 경계가 점점 없어져 가는 경향이 있기 때문에 타학문에 대한 지식이나 이해가 보다 나은 학문 연구나 사회발전을 위해서 필수적이라 할 수 있다.

이와 같은 관점에서 본 교과과정의 개발을 통하여 수리과학이 산업분야에서 커다란 역할을 할 수있음을 보여주는 계기가 됨은 물론 학계 간 연구와 이해가 활발하게 진행 될 수 있는 밑거름이 되기를 기대한다.

## 참고문헌


[1] 박민용,최 항식(1990), *휘지 시스템의 응용입문*

[2] 박용석, 최항식, (1990), *퍼지시스템의 응용입문*, 대영사.

[3] 안성헌, 이주성, 김혜중, 이영섭 (2003). 확률암호시스템 및 응용 교과목개발에 대한 연구, *Journal of the Korean Data Analysis Society*, Vol.5, No.3, pp.677-686.

[4] 유동선,이교원(1998) *기초퍼지이론*, 교우사

[5] 이광형, 오길록. (1991), *퍼지이론과 응용* I,II, 홍릉과학출판사.

[6] 정이채. (1998). 퍼지과학의 세계, 교우사.

[7] Buckley, J.J. and Hayashi, Y. (1994). *Fuzzy Neural Networks. Fuzzy Sets, Neural Networks and Soft Computing*

[8] Manton, K.G., Woodbury, M.A., and Tolley, H.D. (1994). *Statistical Applications Using Fuzzy Sets*, John Wiley & Sons Inc. N.Y.

[9] Negoita, C.V. and Ralescu, D. (1987). *Simulation, Knowledge-Based Computing, and Fuzzy Statisrics,* Van Nostrand Reinhold Company Inc., N.Y..

[10] Zadeh L.A.(1965), Fuzzy Sets, *Information and Control,* 8, pp.338-353.

[11] Zadeh L.A.(1975), *Calculus of fuzzy restrictions, in L.A. Zadeh et.al.,eds. Fuzzy Sets and Their Applications to Cognitive and Desion Processes,* Academic Press, pp. 1-39.