

주택수요 예측인자 영향도 분석에 의한 상관인자선정

The Correlation Factors on the Analysis of Demand Factors for Apartments

양 승 원* · 박 근 준**

Yang, Seung-Won* & Park, Keun-Joon

요 약

주택수요는 수요유발 상관인자를 주축으로 한 주택수요 예측모형에 의하여 그 수요량을 예측할 수 있다. 이때, 주택수요 예측모형은 상관인자의 영향도에 따라서 인자의 미세한 추이변화에 대해서도 수요의 변화 폭을 민감하게 제시하게 된다. 이를 위하여 주택수요 예측에 동원 될 수 있는 여러 상관인자들 가운데 영향도가 가장 큰 인자가 무엇인지 찾아낼 필요가 있다.

이때 대상인자의 데이터는 횡단면자료(Cross Section Data) 혹은 시계열자료(Time Series Data)분석으로 수행된다. 즉, 영향도가 가장 큰 인자들을 찾아내는 방법마련이 필요하며 이후 이 방법에 따른 상관인자의 도출이 가능함에 따라 영향도가 가장 큰 인자를 발굴하는 방법을 제시하고 이에 의한 상관인자를 도출하는 것을 본 연구의 목적으로 한다.

키워드 : 민감도, 상관계수, 상관성, 통계자료, 주성분분석

1. 서 론

1.1 연구의 배경 및 목적

주택의 수요예측은 다소 이론적이고 경험적인 시각에서 연구되어 왔으며 연구주제와 가정에 따라 다소 차이가 있으나 기존 연구 및 문헌을 고찰하여보면 대체적으로 개별 가구주체의 주택 소비를 위한 수요유발인자와 환경 및 정책적지표인자 등이 거론되고 있다.

주택의 수요예측은 수요를 유발하는 인자와의 상관성을 규명한 후 이를 정형화된 모형으로 구축하여 수요예측을 실시할 수 있다. 모형에 사용되는 수요 예측인자는 기타 상관인자 보다도 영향도가 큰 인자들일 때 인자의 미세한 추이변화에 대해서도 수요의 변화 폭을 민감하게 제시하게 된다.

이를 위하여 주택수요 예측에 동원 될 수 있는 여러 상관인자들 가운데 영향도가 가장 큰 인자가 무엇인지 찾아낼 필요가 있다.

따라서, 본 연구는 주택수요 예측인자 가운데 영향도가 가장 큰 인자를 발굴하는 방법을 제시하고 이에 의한 상관인자를 도출하는 것을 연구의 목적으로 한다. 이 때 분석 대상인자의 데이

터는 횡단면 자료(cross section data) 혹은 시계열자료(time series data) 분석으로 수행된다.

인자별 영향도는 추세변화분석에 의하여 영향도를 도출한다.

여기서, 주택수요 상관인자는 수요와의 상관성이 존재하면서 동시에 인자 상호간에도 상관성이 존재할 수 있는 바 이를 규명하여 인자의 우선순위도 결정하여야 한다.

1.2 연구의 범위 및 방법

주택수요 예측인자별 영향도를 도출하기 위하여 우선 영향도 파악 대상인자를 선정하도록 한다. 이를 위하여 기존연구에서 다른 인자들 가운데 정량적 자료취득이 가능한 인자를 영향도 파악 대상인자로 한다.

영향도 분석 대상인자는 피어슨의 상관성이론에 의한 상관계수 크기가 확보된 인자로 한다. 상관계수 크기가 확보된 상관인자의 영향도 분석방법으로서 단순회귀분석과 중회귀분석을 실시한다.

한편, 추출된 인자들은 상관계수행렬(correlation matrix)을 이용 상호관계를 분석하고 인자간 인과관계 다이어그램(diagram)에 의한 인자들의 상호 독립성이 파악된 인자들을 분석대상으로 한다.

본 연구의 주택수요 대상은 단독주택을 제외한 아파트 수량으로 한다.

* 일반회원, 호서대학교 대학원 박사과정

** 종신회원, 호서대학교 건축공학과 교수, 공학박사

본 연구는 호서대학교 교내연구비 지원에 의한 것입니다.

2. 주택수요예측 영향인자

2.1 주택수요개념

수요개념은 일반적으로 재화나 용역에 대한 단순한 소유 욕구가 아닌 구매력(購買力)이 수반된 욕구로서, 경제적 부담능력을 갖춘 상태에서 어느 정도의 가격에서 얼마만큼의 양을 소비할 것인가를 나타내는 것이다¹⁾.

주택수요(housing demands)는 주택 소비에 대한 수요를 말하는 것으로 주택 유효수요²⁾를 의미하는데 이는 주거서비스를 제공하는 필수재로서 주택에 대한 수요(실수요)와 재산축적 수단으로서의 주택자산(투기적수요)에 대한 수요이다.

주택건설업체 입장에서 주택수요는 주택건설사업 타당성에 입각하여 추정된 적정 주택공급 규모이다. 이는 주택 유효수요 규모로서 주택 초과공급으로 인한 부작용인 미분양을 방지³⁾하기 위해 추정된 주택수요량이다.

2.2 주택수요예측 영향인자

산업적 측면에서 적정주택공급의 의미⁴⁾는 지역주택공급량 결정을 위한 주택유효수요로 할 수 있다. 주택수요예측을 위해 적용되는 방법은 추정모형구축⁵⁾을 통한 수요예측인데 수요추정에 이용되는 예측인자나 설명변수들이 주택수요예측 영향인자로서 설정된다.

우선, 주택수요예측 영향인자들을 추출하기 위한 기존 연구 및 문헌에 의하면 김혜선(1996)은 주택문제 해결을 위해 주택수요(needs) 및 수요(demands)의 이론적 고찰을 하면서 주택소비를 일으키는 것을 주택수요라 정의하고 주택수요 영향인자를 주택소비를 위한 가구의 주거이동요인이라 정의하였다. 따라서, 주거이동에 영향을 미치는 요인을 주택수요 영향인자로서 제시하였는데 이는 주택가격(housing price), 가계소득(family income), 가족수 및 가족구성, 투자가치, 자가소유욕구, 선호(preference), 가정생활주기(family life cycle : FLC) 등이다.

1) 주택수요 및 주택수요에 관한 고찰, 1996, 김혜선
 2) 경제학에서의 수요는 보통 유효수요로서 지칭하나 유효수요란 '구매능력'과 '구매의사'를 갖춘 수요이다.
 3) 90년대 초반 적정규모를 초과한 과대공급으로 인해 주택 미분양 추세가 증가하곤 하였다. 주택 미분양은 주택업체의 자금부담을 심화시키고, 사회적으로도 자원의 비효율성을 초래한다. 따라서 주택미분양 방지를 위한 관점에서 적정주택공급량은 주택수요량으로서 주택시장에서 최대한으로 소화 가능한 주택공급 규모이다.
 4) 관련 선행연구를 종합해 볼 때 국제 비교적 관점, 자산균형모형, 잠재생산력, 정책적, 산업적 측면의 적정주택공급을 그 의미의 유형이라 할 수 있다.
 5) 대체적으로 횡단면자료(Cross Section Data) 혹은 시계열자료(Time Series Data)에 의한 평활법, 회귀분석방법, 박스-젠킨스 분석으로서 구축된 추정모형이다.

표 1. 주택시장분석의 분류 및 분석내용⁶⁾

분류	분석요인	분석내용
거시적 분석	수요영향 요인	인구, 연령구조, 가구규모, 성별, 고용, 관습 및 선호도, 가격상승, 입지경쟁
	공급영향 요인	건축착공량, 건축허가량, 공실률, 임대료추세, 이용가능성, 가격, 도시·지역계획, 건축비용, 금융유동성
미시적 분석	인근분석요인	지역경제영향, 토지이용상황, 지역명성, 지역평가, 경쟁지역 분양가·임대가, 인근 교통망, 편의시설 접근성, 용도지역검토, 개발가능면적, 편의시설 접근성, 지역내 입지경쟁시설, 공실률추세, 근접 편의시설 효과, 장래 인근지역 발전추세
	대지 및 건물분석	요인 면적, 형상, 용도지역, 지질, 상하수도, 편의시설 접근거리 및 시간, 경쟁대지와의 비교, 용도지역지구

표 2. 주택수요예측 영향인자

No.	저자	년도	연구주제	수요예측 영향인자
1	김혜선	1996	주택수요 및 수요에 관한 고찰	주택가격, 가계소득, 가족수 및 구성, 투자가치, 자가소유욕구, 선호, 가정생활주기
2	오동훈	1998	수요측면에서 본 광역주택시장분석	소득 및 지출, 주거소비수준, 주택구매력, 주거만족도
3	류삼열 외	1998	인구규모에 따른 주택공급여건분석	인구, 가구, 주택, 기반시설, 토지이용, 산업
4	허재완	2000	지역별 적정 주택공급규모 예측	전년도 주택수요, 총인구, 소득, 주택가격, 이자율, 더미변수(IMF)
5	김양렬	2001	주택수요예측을 위한 생태학적 모형	총인구, 주택보급율, 평균가구원수, 소득, 연령 구조
6	안정근외	2003	부동산학개론	표1 참고

오동훈(1998)은 개발수요 추정시 선행되어야 할 지역 주택시장 분석기법의 기본틀을 제시하면서 지역주택 시장분석을 위한 분석인자를 지역소득 및 지출, 주거소비수준, 주택구매력, 주거만족도 등으로 설정하고 광주광역시 주택시장을 분석하였다.

류삼열 외(1998)는 지역특성과 인구규모에 따른 분양성 예측모형을 제시하면서 관련 분석인자를 인구, 가구, 주택, 기반시설, 토지이용, 산업 등으로 선정하고 지역별로 다른 분양률 예측함수를 제시하였다.

허재완(2000)은 지역별 적정 주택공급규모를 추정하기 위한 유효수요 추정모형 구축을 위해 시계열자료(time series data) 분석에 의한 회귀모형을 구축하고 주택유효수요 규모를 추정하였다. 모형구축에 사용된 파라미터의 설명변수(explanatory variable)는 지역별로 해당하는 전년도 주택수요, 총인구, 소득, 주택가격, 이자율, 더미변수(IMF 금융대란)등이다.

김양렬(2001)은 주택수요량을 예측하기 위한 회귀모형구축을 하면서 설명변수를 총인구, 주택보급율, 평균가구원수, 가구소득, 연령구조 등으로 설정하고 주택수요량을 추정하였다.

6) 안정근 외, 부동산학개론, 법문사, 2003, p.122

안정근(2003)은 주택시장을 거시적 분석을 위한 수요 및 공급 영향요인, 미시적 분석을 위한 인근지역과 대지 및 건물분석요인으로 분류하여 제시하였다.

기존연구 및 문헌을 고찰한 결과 주택수요예측 영향인자는 연구자의 다양한 관점과 자의적인 해석에 따라 추출기준 및 분석 목적들이 상이하여 다소 차이가 있으나 어느 지역에서나 보편적으로 적용 가능하고 기존연구에서 적용한 빈도수를 고려할 때 인구, 가구, 경제, 산업시설, 생활기반시설, 주택가격, 소득, 가족수 등으로 정리되고 있다⁷⁾.

3. 주택수요예측을 위한 영향인자 영향도분석

3.1 분석대상 인자추출 및 상관성 분석

주택수요 예측인자 영향도 분석을 위한 전단계로 우선 상관성 분석대상 인자를 추출하도록 한다.

분석대상 인자를 표2의 기존연구 및 문헌⁸⁾에 의한 통계치로서 대상인자를 추출토록 하였으나 통계자료취득이 용이하지 못한 인자는 정량적 데이터로서 객관적인 분석이 불가능하기 때문에 추출대상 인자범위에서 제외한다.

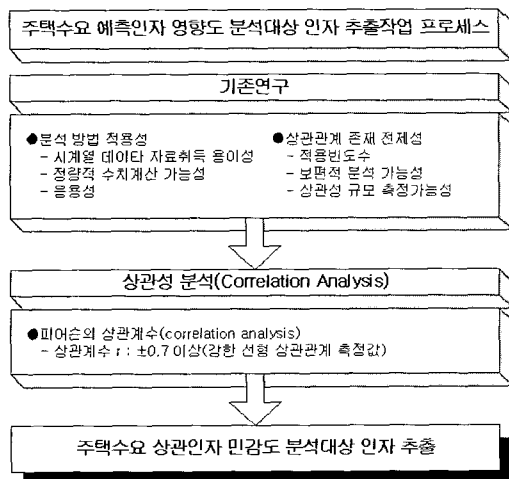


그림 1. 분석대상 영향인자 추출작업 프로세스

그림 1은 주택수요와의 상관성을 분석할 대상인자를 추출하고 상관성이 검증된 인자를 대상으로 영향도를 분석하는 프로세스이다.

영향도 분석 대상인자는 정량적 데이터로서 피어슨의 상관성 이론⁹⁾에 의한 상관계수(correlation coefficient) 크기가 확보된

인자이다.

표 3. 상관성 분석대상 설명변수

No.	설명변수(Explanatory Variable) 및 측정방법	
1	X1	인구수 현황(인)
2	X2	가구수 현황(가구)
3	X3	가구원수(인, 인구수/가구수)
4	X4	주택보급율(%)
5	X5	단독주택비율(%), (단독주택수/총주택수)×100)
6	X6	도로율(%), (폭4m이상도로면적/시가지면적)×100)
7	X7	포장율(%)
8	X8	상수도보급율(%)
9	X9	승용차자가용수(대)
10	X10	주거용지비율(%), (주거용지면적/도시계획면적)/100)
11	X11	교육시설수(초·중·고 학습수, 대학이상·학과수)
12	X12	1인당지방세부담액(원)
13	X13	소비자물가지수(2000년 기준 연평균지수)
14	X14	제조업체수(업체수)

근거자료: 1. 천안시 통계연보
2. 내무부 한국도시연감

적용성이 가능한 인자들인 설명변수(explanatory variable)를 나열하면 다음 표3과 같다.

표3의 분석대상인자들은 주택수요 예측인자로서 인구, 가구, 가구원수 경제, 산업시설, 생활기반시설, 주택가격, 소득 등을 표현하고 설명할 수 있는 통계지표로서 종속변수에 대한 설명변수¹⁰⁾이며 선형성 있는 주택수요 예측모형에 이용되는 상관성 분석 대상인자이다.

상관관계를 분석하기위해 주택수요를 의미하는 종속변수¹¹⁾는 연별 주택량으로 설정하였다.

주택수요와 예측인자간의 개별적인 상관관계는 두 양적 변수간의 선형관계하에 분석 가능하다. 따라서, 종속변수인 공동주택의 연도별 세대수 증가추세가 어떠한 추세로서 분포성향을 나타내고 있으며 상관관계분석 전제조건으로서 선형성을 만족시키는지 여부를 우선적으로 파악하고자 분포추이 적합곡선을 검토한다.

곡선추정(curve estimation)은 자료 분포형태가 어떤 추세를 나타내며 도출된 추세를 바탕으로 미래 예측구간을 결정함에 유용하다. 또한 인과관계 없이 자체 자료에 대한 최적합 함수형태를 추정하기 위한 것으로 예측함수구축을 위한 다양한 곡선조건을 선별할 수 있다. 이는 해당 자료에 대한 표4의 통계 결과를

7) 박근준 외, 천안시 아파트 수요추정을 위한 모형구축방법, 대한건축학회 논문집 제19권 11호, 2003. 11., p193

9) Person's simple correlation coefficient

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(n-1) S_x S_y}$$

10) 주택수요 증감현상에 상관관계가 예상되는 양적변수로서의 독립변수(independent variable)이며 천안시 1990~2001년 통계자료를 이용하였다.

11) 본 연구는 주택수요를 의미하는 종속변수(dependent variable)를 천안 지역 연별 주택량 현황(세대수, 양적변수) 추이로서 설정하였고 천안시 1990~2001년 통계자료를 이용하였다.

근거로 가능하다.

천안지역 공동주택수량 증감현상이 어떠한 곡선형태를 나타내고 있고 선형성을 검토하기 위해 선형함수 적합성을 검토하고자 SPSS 곡선추정 프로그램을 이용하여 종속변수의 시간흐름(1990~2001)에 따른 양적 증가추세와 이에 적합한 곡선을 작성하면 다음 그림 2와 같다.

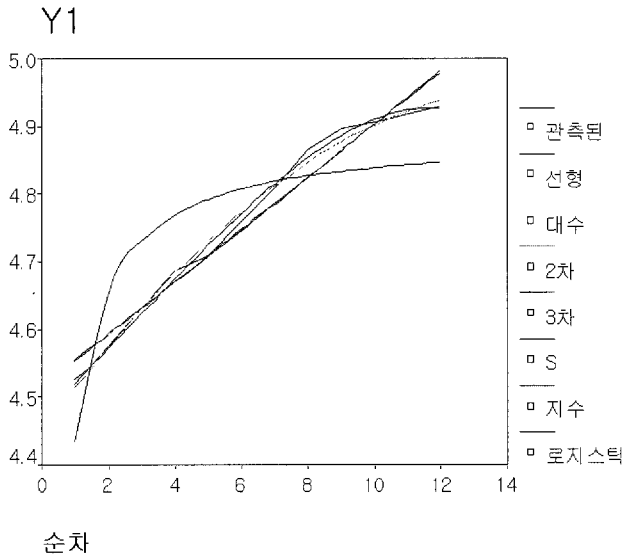


그림 2. 종속변수 곡선추정

표 4. 추정곡선 함수별 분석결과

Dependent	Mth	Rsqr	d.f.	F	Sigf	b0	b1	b2	b3
Y1	LIN	.963	10	257.36	.00	4.515	.038	-	-
Y1	LOG	.950	10	190.81	.00	4.461	.183	-	-
Y1	QUA	.993	9	622.27	.00	4.447	.067	-.002	-
Y1	CUB	.996	8	642.88	.00	4.480	.042	.002	-.0002
Y1	S	.722	10	25.99	.00	1.58	-.096	-	-
Y1	EXP	.960	10	237.38	.00	4.51	.008	-	-
Y1	LGS	.960	10	237.38	.00	.221	.991	-	-

Independent : Time

종속변수의 곡선추정 결과 S-모형곡선을 제외한 대부분의 모형곡선이 높은 R² 값을 나타내고 있다. 따라서 S-모형 곡선을 제외한 나머지 추정곡선적용이 가능하나 선형을 전제로한 선형함수 곡선인 LIN을 선정한다.

선형함수의 곡선인 LIN의 R² 값을 살펴보면 0.96으로서 종속변수를 선형곡선에 적합시킬 경우 기존 분포데이터에 대한 LIN의 설명도 정도가 96%임을 의미한다. 또한 종속변수인 천안지역 공동주택수 증가량의 추세를 선형적 증가추세로 예측할 경우 회귀계수 0.38, 오차 4.52인 자체 적합곡선 추정 파라미터로서

제시된다.

한편, 종속변수와 상관성 있는 각각의 예측인자들이 천안지역 주택수요와 어떠한 형태로 개별적인 상관관계를 나타내고 있는지 파악키 위해 산점도를 작성하였다. 이는 선형성이 검증된 종속변수에 대한 설명변수별 데이터 분포성향을 파악하기 위함이다.

작성된 산점도상의 데이터 분포형태는 변수간의 강하고 약한 정도 및 음의관계, 무상관관계정도를 가늠 할 수 있도록 하고 있다. 여기서 Y축은 주택수요량, X축은 개별 예측인자이다.

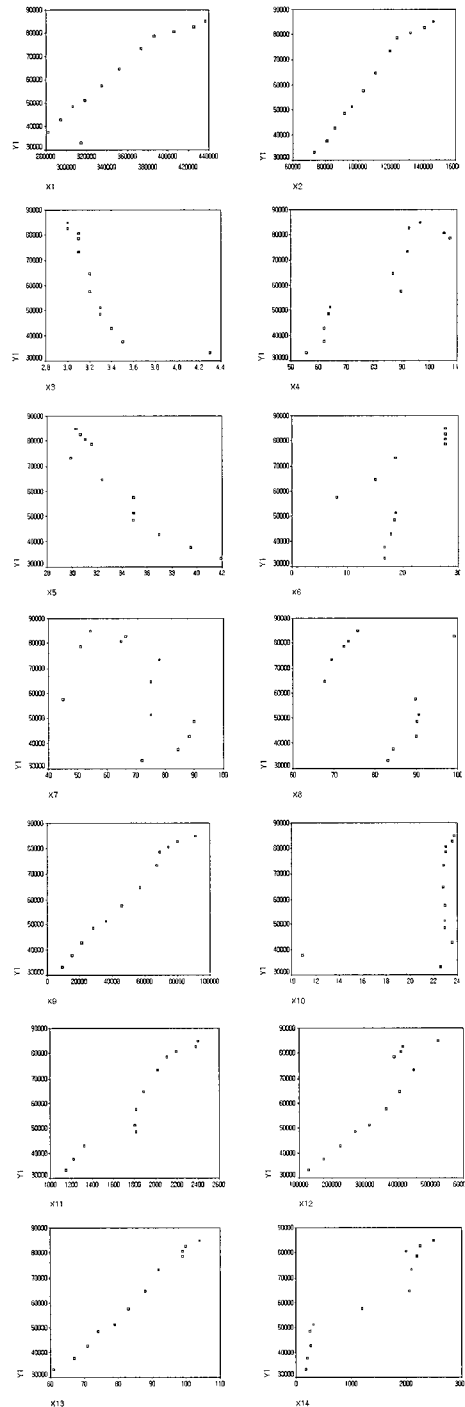


그림 3. 설명변수별 산점도

12) 회귀변동이 총 변동중에서 차지하는 비율로서 결정계수(coefficient of determination)이다.

SPSS에 의한 그림 3의 산점도들은 대체적으로 천안지역 주택 수요와 설명변수간의 상관관계가 설정되고 있다.

그림 3의 산점도중에서 X1, 2, 3, 9, 13의 산점도 분포가 강한 선형성을 띠고 있으며 X6, 8, 10의 산점도는 다소 선형성이 부족하기 때문에 인자변환의 필요성이 요구되고 있다.

산점도의 데이터 분포와는 별도로 상관관계 정도를 측정하고자 피어슨의 상관계수(correlation coefficient)에 의한 상관성 크기를 계산하였다.

이를 위해 SPSS 프로그램에 의한 상관계수를 측정한 결과는 표5와 같이 요약되고 있다.

강한 선형 상관관계를 의미하는 상관계수 r의 크기 범위는 ± 0.7 이상¹³⁾으로 표3의 설명변수별 상관계수 크기를 살펴보면 X6, 7, 8, 10 등을 제외한 나머지 주택수요 설명변수들이 천안지역 주택수요와 매우 강한 상관관계가 존재하고 있음을 알 수 있다.

예측인자별 측정된 상관계수와 유의확률을 검토하면 우선 X7(포장율)과 X8(상수도보급율)의 상관계수가 -0.57, -0.36으로서 다소 상관성이 부족하며 유의도 또한 통계적 유의도 기준인 0.05보다 큰 0.052, 0.244로 나타나고 있어 통계적으로 유의하지 못하다. 또한 X10(주거용지비율)의 상관계수 및 유의도가 0.44, 0.15로서 다소 상관성이 부족하며 유의하지 못하다. X6(도로율) 역시 상관성이 부족하게 나타나고 있다. 이는 천안지역 공동주택 수요가 도시기반시설인 X6(도로율), X7(포장율), X8(상수도보급율), X10(주거용지비율)과는 별로 상관성이 없음을 설명한다. 즉, 천안지역 주택수요가 일정하게 증가하는 추세와는 달리 도시기반시설과 주거용지비율의 증가추세가 일치하지 않고 있음을 의미한다.

이것은 소규모 도시에 한해서 주택공급량 증가와 도시기반시설 및 주거용지비율과 밀접한 상관인자로서 분석된 기존연구¹⁴⁾와는 다소 다른 결과를 보이는 것이다.

상관관계를 전제로 영향도 분석대상 인자들은 X1, 2, 3, 4, 5, 9, 11, 12, 13, 14 들로서 회귀분석에 의한 회귀계수크기 영향도 분석을 수행할 수 있다.

3.2 단순회귀분석

상관성이 검증된 상관인자 각각의 영향도를 분석하기 위한 방법으로 우선 단순회귀분석에 의한 회귀계수 도출을 수행한다.

단순회귀 분석결과에서 도출가능한 회귀계수는 주택수요 상관인자별 영향도를 분석하기 위한 수요에 대한 탄력성을 의미한다. 즉, 단순회귀분석에 의한 회귀계수의 크기로 탄력성이 크고 작음을 파악할 수 있다. 여기서 상관성이 검증된 인자들을 대상으로 단순회귀모형 구축이 가능하다.

회귀분석을 위해 SPSS프로그램을 이용하여 회귀계수를 추출하고자 우선 종속, 설명변수 값에 log값을 취하여 변수 전체 척도들을 통일하였다.

표6에 의하여 회귀계수 크기 순서가 통계적 유의도에 별다른 문제가 없는 내에서 양으로서 X1, 13, 2 등의 순으로, 음으로서 X3, 5 등의 순으로 측정되고 있다. 특히 X1, 3은 인구수와 가구원수로서 천안지역 인구수가 증가하고 세대별 가구원수가 감소하는 현상이 주택수요와 양의관계, 음의관계로서 각각 계수가 2.03, -2.97로서 가장 영향력 있는 상관인자 로서 제시되고 있다.

표 6. 설명변수별 단순회귀분석(탄력성:회귀계수)결과

NO	EV	B	Beta	R2	t - α	F - α
1	X1	2,038	0,923	0,852	7,598	0
2	X2	1,43	0,989	0,978	20,992	0
3	X3	-2,978	-0,89	0,79	-6,141	0
4	X4	1,295	0,941	0,886	8,8	0
5	X5	-2,915	-0,97	0,946	-13,18	0
6	X9	0,439	0,986	0,972	18,599	0
7	X11	1,262	0,962	0,925	11,094	0
8	X12	0,739	0,95	0,903	9,67	0
9	X13	1,863	0,997	0,994	41,508	0
10	X14	0,278	0,949	0,901	9,567	0

범례- EV: 설명변수(Explanatory Variable)
 B: 회귀계수
 beta: 표준회계수
 t-α: t값 - 유의수준
 F-α: F값 - 유의수준

표 5 설명변수별 Pearson 상관계수

EV	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14
r	.952**	.988**	-.801**	.932**	-.944**	.684*	-.571	-.365	.993**	.441	.958**	.937**	.996**	.952**
유의확률	.000	.000	.002	.000	.000	.014	.052	.244	.000	.151	.000	.000	.000	.000
N	12	12	12	12	12	12	12	12	12	12	12	12	12	12

** 상관계수는 0.01 수준(양쪽)에서 유의
 * 상관계수는 0.05 수준(양쪽)에서 유의

13) 윤태연, 정성원 공저, SPSS 아카데미 통계조사분석, 고려정보산업, 1999, P.297

14) 류삼열 외, 인구규모에 따른 주택공급여건분석, 1998

표 7. 영향도 순위

순위	EV	회귀계수	R ²
1	X1	2.038	0.852
2	X13	1.863	0.994
3	X2	1.43	0.978
4	X4	1.295	0.886
5	X11	1.262	0.925
6	X120	.739	0.903
7	X9	0.439	0.972
8	X14	0.278	0.901
1	X5	-2.978	0.79
2	X3	-2.915	0.946

천안지역 소비자 물가지수(2000년기준 연평균지수)와 가구수를 의미하는 X13, 2도 천안지역 주택수요에 민감한 상관인자로서 제시되고 있다.

X5인 단독주택비율은 음의 값으로 공동주택 패턴과 반대방향으로 움직이고 있다. 이는 주택수요구조의 변화를 의미한다.

단순회귀분석 결과에 따라 천안지역 주택수요 우선 분석대상 인자는 X1, 5 등으로 제시되고 있다.

3.3 인자분석에 따른 중회귀분석

상관성이 검증된 인자들을 대상으로 인자축소를 위하여 상관행렬(correlation matrix)에 의한 전체 인자들을 설명하고 대표할 수 있는 공통인자를 추출한다.

공통인자 추출은 기존 상관인자들의 특성을 유지하면서 대표인자로서 영향도 분석이 가능하도록 한다.

공통인자 추출은 인자들의 상관행렬 및 공분산행렬의 고유값과 고유벡터를 이용한 스펙트럴 분해(spectral decomposition)를 이용하여 주성분이라는 새로운 변수를 생성하는 방법인¹⁵⁾ 주성분 분석(principal component analysis)으로 가능하다. 따라서, 상관성이 검증된 인자들을 대상으로 상관행렬 및 주성분분석을 수행하여 공통인자를 추출하고 추출된 공통인자를 대상으로 중회귀분석에 의한 표준화계수(Beta)를 측정한다. 측정된 표준화계수는 종속변수를 설명하기 위한 상대적인 중요도를 의미하며 이는 회귀모형에서 종속변수와 독립변수들을 평균이 0, 분산을 동일하게 표준화한 후 얻어진 회귀계수이다.

우선, 인자간의 상관행렬(correlation matrix)을 분석하면 추출할 공통인자의 개수를 결정할 수 있다. 인자별 상관관계를 표 8의 상관행렬을 기준으로 분석하면 인자별 상관관계가 다양하

나 X2와 X1, 4, 5, 9, 11, 13, 14와 상관계수 값이 크고 유사하며 X3와 X12가 밀접한 상관관계로서 유사한 집단을 형성하기에 두 집단으로서 공통인자 제시가 가능하다. 이에 따라, 두개의 공통인자 추출을 위한 주성분 분석을 수행한 결과 표9와 같이 고유값 1을 기준으로 2개의 주성분 공통인자가 보유되고 이들의 누적기여율은 96.3%로 나타나고 있다. 그림 5의 스크리 도표는 고유값에 따른 크기의 변화를 나타내는데 3번째공통인자부터는 기울기의 변화가 크지 않으므로 2개의 공통인자를 선정하는 것이 바람직하다.

표 8. 상관행렬(correlation matrix)

	X1	X2	X3	X4	X5	X9	X11	X12	X13	X14
X1	1.000	.948	-.678	.867	-.852	.864	.874	.803	.922	.898
X2	.948	1.000	-.877	.913	-.949	.969	.954	.928	.990	.930
X3	-.678	-.877	1.000	-.798	.907	-.933	-.883	-.932	-.893	-.793
X4	.867	.913	-.798	1.000	-.873	.915	.849	.869	.938	.966
X5	-.852	-.949	.907	-.873	1.000	-.976	-.950	-.964	-.961	-.904
X9	.864	.969	-.933	.915	-.976	1.000	.967	.986	.986	.932
X11	.874	.954	-.883	.849	-.950	.9671	.000	.949	.957	.862
X12	.803	.928	-.932	.869	-.964	.986	.9491	.000	.950	.902
X13	.922	.990	-.893	.938	-.961	.986	.957	.9501	.000	.945
X14	.898	.930	-.793	.966	-.904	.932	.862	.902	.945	1.000

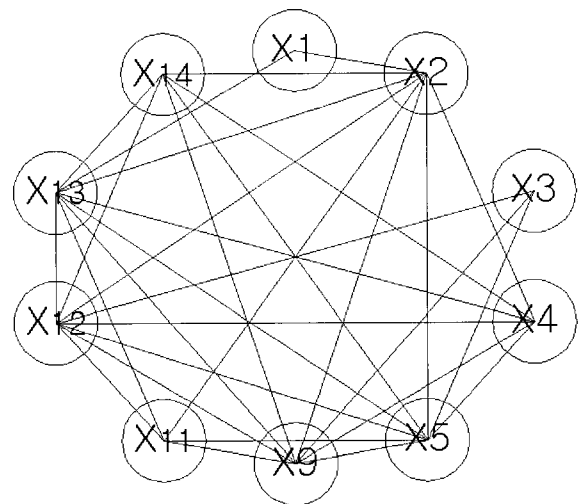


그림 4. 인자별 상관관계 다이어그램(diagram)

주성분 분석에 의한 표10의 인자적재행렬을 살펴보면 공통인자인 Fac-1(성분1)은 X3, 5가 음의계수로서 나머지 인자는 모두 양의계수로 제시된다. Fac-2(성분2)는 X1과 3을 제외하고는 모든 인자가 낮은 계수로서 제시되고 있다. 공통인자별 용이한 해석을 위해 Varimax 방법에 의한 인자회전을 실시한 후 회전공간내 인자 분포현황을 도출하면 그림 6과 같다. 회전공간내 인자분포 현황을 살펴보면 Fac-1(성분1)은 X1, 2, 4, 9, 11, 12,

15) 박성현 외 공저, SPSS 아카데미 P.377 ~ 395

13, 14를 설명하는 대표인자이며 Fac-2(성분2)는 X3, 5를 설명하는 대표인자라 할 수 있다. Fac-1과 2의 세부 설명인자들의 가장 큰 특성은 음의 값과 양의 값을 갖고 있는 점이다. 즉, 데이터가 증가와 감소하는 분포현상으로 그림 7과 같이 분류된다.

표 9. 주성분 분석

성분	초기 고유값			추출제곱합 적재값			회전제곱합 적재값		
	전체	%분산	%누적	전체	%분산	%누적	전체	%분산	%누적
1	9.201	92.005	92.005	9.201	92.005	92.005	4.815	48.154	48.154
2	.425	4.251	96.256	.425	4.251	96.256	4.810	48.102	96.256
3	.199	1.990	98.246						
4	7.732E-02	.773	99.019						
5	4.863E-02	.486	99.505						
6	3.236E-02	.324	99.829						
7	1.084E-02	.108	99.937						
8	5.616E-03	5.616E-02	99.993						
9	6.704E-04	6.704E-03	100.000						
10	3.521E-06	3.521E-05	100.000						

추출방법 : 주성분분석

표 10. 인자적재행렬

	성분	
	1	2
X1	.908	.367
X2	.986	7.414E-02
X3	-.907	.383
X4	.937	.189
X5	-.974	.107
X9	.994	-9.46E-02
X11	.964	-8.43E-02
X12	.969	-.187
X13	.995	3.842E-02
X14	.952	.196

요인추출 방법 : 주성분 분석

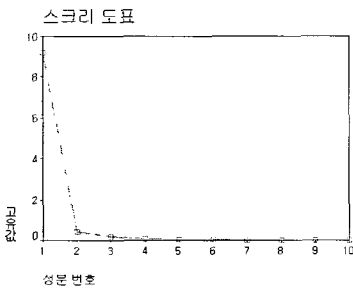


그림 5. 스크리도표

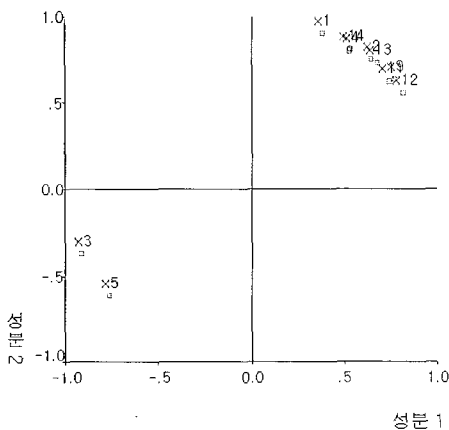


그림 6. 회전공간내 인자분포현황

Fac-1의 세부인자들은 천안지역 인구수와 상호관련이 있는 인자와 지역경제 및 산업시설을 의미하는 인자들의 집단으로서 지역인구 및 경제관련 공통인자로서 제시되며 Fac-2의 세부인자는 개별가구의 가구원수 감소 및 단독주택비율 감소현상을 설

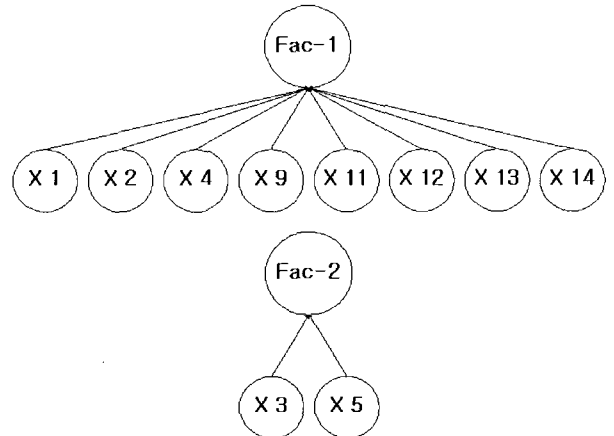


그림 7. 공통인자별 세부인자 분포현황

명하는데 이는 가구원수 감소에 따른 주택공급형태 변화를 의미하는 공통인자로서 제시된다.

표11의 Fac-1,2의 데이터는 주성분 분석 후 제시되는 공통인자의 추이로서 이를 대상으로 종속변수와 중회귀분석을 수행한다.

표 11. 공통인자(성분1,2)의 데이터(Log)

	Fac-1	Fac-2
1	-2.96045	0.35995
2	-0.59778	-1.26193
3	-0.12452	-1.23611
4	0.57045	-1.37091
5	0.64328	-1.16986
6	0.14706	-0.02832
7	0.36932	0.18233
8	0.62372	0.36244
9	0.11546	0.98991
10	0.22029	1.0614
11	0.39277	1.02557
12	0.60041	1.08553

Fac-1,2를 대상으로 종속변수에 대한 단계선택법(stepwise selection)으로 중회귀분석(multiple regression analysis) 후 도출되는 표준화계수(standardized regression coefficient)는 종속변수에 대한 상대적 중요도를 판단하는 지표¹⁶⁾이다. 중회귀 분석에 의한 표준화계수 추출결과는 표12와 같다.

표 12. 성분1, 2 중회귀분석 표준화계수 추출결과

모형	비표준화 계수		표준화계수		유의확률
	B	표준오차	베타	t	
1 (상수)	4.767	.029		163.022	.000
REGR factor score 2 for analysis 1	.104	.003		1627.078	.000
2 (상수)	4.767	.003		1627.078	.000
REGR factor score 2 for analysis 1	.104	.003	.734	34.087	.000
REGR factor score 1 for analysis 1	9.615E-02	.003	.676	31.419	.000

a. 종속변수:Y1

16) 김두섭 외 공저, 회귀분석, 나남출판 p.59~61

모형	분산분석						R ²
	선형	제공합	자유도	평균제곱	F	유의확률	
2	선형	.221	2	.111	1074.542	.000	.996
	잔차	.103	9	1.030E-04	-	-	-
	합계	.222	11	-	-	-	-

공통인자별로 추출된 표준화계수 크기를 살펴보면 Fac-2의 표준화계수 크기가 Fac-1보다 0.6정도 큰 베타값으로 측정되고 있어 Fac-2의 소속인자가 Fac-1보다 종속변수인 공동주택수요에 대해 상대적으로 더 중요한 것으로 판단되기는 하나 공통인자 0.6정도의 표준화계수 크기로서 중요도를 판정하기에는 적은 수치이기 때문에 중회귀분석에 의한 공통인자에 소속된 세부인자별 표준화계수를 비교하기로 한다.

우선 Fac-1의 세부 소속인자 중회귀분석 후 도출된 표준화계수는 다음 표13과 같다.

단계선택법에 의한 중회귀분석 결과 X13이 단독으로 진입되어 표준화계수 비교분석이 불가능함에 따라 X13을 제외한 나머지 인자로서 표준화계수를 도출하였다.

표13의 중회귀분석에 의한 표준화계수 추출결과, X2, 4, 9, 4를 제외한 나머지 인자들은 선형을 가정한 중회귀분석이 통계적 유의도결여로 계수추출에서 제외되었다.

표 13. Fac-1 세부인자 중회귀분석 표준화계수 추출결과

모형	비표준화 계수				표준화계수	t	유의확률
	B	표준오차	베타	t			
1 (상수)	-2.420	.342		-7.066	.000		
	1.430	.068	.989	20.992	.000		
2 (상수)	-1.155	.694		-2.23	.828		
	.794	.189	.549	4.201	.002		
	.202	.058	.454	3.470	.007		
3 (상수)	8.631E-02	.570		.151	.883		
	.695	.158	.481	4.392	.002		
	.165	.049	.372	3.344	.010		
	.223	.093	.162	2.400	.043		

a. 종속변수:Y1

모형	비표준화 계수				표준화계수	t	유의확률
	B	표준오차	베타	t			
1 (상수)	1.187	.086		13.750	.000		
X13	1.863	.045	.997	41.5508	.000		

a. 종속변수:Y1

모형	분산분석						R ²
	선형	제공합	자유도	평균제곱	F	유의확률	
3	선형	.221	3	7.370E-02	480.622	.000	.994
	잔차	1.227E-03	8	1.533E-04	-	-	-
	합계	.222	11	-	-	-	-

모형	분산분석						R ²
	선형	제공합	자유도	평균제곱	F	유의확률	
1	선형	.221	1	.221	1722.900	.000	.994
	잔차	1.283E-03	10	1.283E-04	-	-	-
	합계	.222	11	-	-	-	-

분석결과, X2, 9, 4의 순으로 표준화계수 크기를 비교할 수 있으며 이는 개별 회귀계수 추출시 X1(인구수)가 가장 큰 계수가 되는 것과 달리 가구수, 승용차자가용수, 주택보급율이 주요한 상관인자가 되고 있다. 즉, Fac-1의 세부인자의 주택수요에 대한 상대적 중요도를 분석한 결과, 가구수, 승용차자가용수, 정책적 지표라 할 수 있는 주택보급율이 주요한 인자가 되고 있다.

Fac-1과 마찬가지로 Fac-2의 세부인자별 중회귀분석에 의한 표준화계수를 추출한 결과는 다음 표14와 같다.

Fac-2의 소속인자인 X3, 5의 표준화 계수는 각각 -0.039, -0.937로서 종속변수에 대해 X3보다 X5가 상대적으로 더 큰 중요도를 나타내고 있다.

표 14. Fac-2 세부인자 중회귀분석 표준화계수 추출결과

모형	비표준화 계수		표준화계수		t	유의확률
	B	표준오차	베타	t		
1 (상수)	9.133	.572		15.954	.000	
X3	-.129	.618	-.039	-209	.000	
X5	-2.810	.553	-.937	-5.082	.001	

a. 종속변수:Y1

모형	분산분석						R ²
	선형	제공합	자유도	평균제곱	F	유의확률	
3	선형	.222	2	.105	78.584	.000	.946
	잔차	1.204E-2	9	1.338E-03	-	-	-
	합계	.222	11	-	-	-	-

이를 해석하면 음의 성향을 나타내고 있는 Fac-2의 세부인자는 X3, 5로서 X3의 세대별 평균 가구원수에 비해 단독주택비용 감소비율이 공동주택 수요에 더 큰 중요도를 나타내고 있다.

4. 결론

주택수요를 유발하는 상관인자를 주축으로 주택수요 예측 모형을 구축 후 이를 이용하여 유효한 주택수요량을 예측할 수 있다.

일반적으로 주택수요를 유발시키는 여러 인자들 가운데 일정한 상관성이 존재하는 인자들이 있는데 이 인자들 가운데 주택수요에 가장 크게 영향을 미치는 인자를 찾아내면 유효주택 수요예측에 비중있게 이용될 수 있다. 이에 따라, 주택수요에 영향을 크게 미치는 인자가 어떠한 것이 있는지 본 연구를 통하여 진행하였다.

주택수요 예측인자 상관관계 및 영향도를 주택수요와 상관성 있는 인자를 시계열 데이터로서 자료취득의 용이성, 정량성 및 응용성등에 의거 통계자료로서 열거하였다. 열거된 예측인자들을 대상으로 주택수요에 대한 산점도를 작성하고 상관계수를 측

정하여 상관관계가 검증된 인자들을 영향도 분석대상 인자로 선정하였다.

상관인자별 영향도 분석결과, 개별인자별 단순회귀 분석에 의하면 주택수요는 인구수 증감현상에 가장 크게 민감한 반응을 보이고 있으며 인자분석에 따른 추출된 공통인자별 세부인자를 중회귀분석한 결과에 의하면 가구수와 단독주택량이 추출된 공통인자 내에서 중요한 지표로 민감한 반응을 보이고 있다.

이는 지역인구 및 가구수가 주택수요를 분석하고 예측하기 위한 가장 중요한 지표임을 의미한다. 이것은 세대별 가구원수 감소 및 독신가구 증가로 인한 가구수 증가는 주택공급형태를 단독주택에서 공동주택으로 전환시키는 작용을 한다고 할 수도 있다.

주택수요를 예측함에 있어 인구 및 가구수 증감은 중요한 분석인자로 제시되는 것이며 이를 증감시키는 영향인자는 상관관계가 있는 정성·정량적 주택수요 영향인자가 된다. 이에 따라, 인구 및 가구수를 증감시키는 다양한 영향인자가 무엇인지를 발굴 분석하는 것은 향후 연구과제로 제시된다.

또한, 본 연구결과는 공동주택수요를 예측하기 위한 수요추정 모형구축에 적용할 수 있을 것이다.

참고문헌

1. Alfred A. Ring & J. Dasso, Real Estate Practices, 1997.
2. Appraisal Institute, The Appraisal of Real Estate, 10th ed., Chicago: Appraisal Institute, 1992.
3. 김두섭 외, 회귀분석, 나남출판, 2000.
4. 박성현, 회귀분석, 민영사, 2001.
5. 박성현 외, SPSS 아카데미, (주)데이터솔루션, 2003.
6. 이덕기, 예측방법의 이해, (주)데이터솔루션, 2002.
7. 이현 외, 주택시장 분석을 위한 자료조사 체계 및 기법 연구, 대한주택공사 주택연구소, 1992.
8. 한국도시연감, 행정자치부, 1990~2001.

Abstract

This research describes an interactive process of analysing the demand factors for apartment on Cheonan area.

Using subjective statistical data for demand factor the process are categorized into main factors explained for the sensitiveness of correlation coefficient. This investigation is based on an analysis of the work of time series data.

One of the propose of this research is determining the correlation factors that can be effectively used in the model of forecasting.

The results show a significant correlation coefficient on correlation matrix to find the optimum correlation factors. The paper thus shows how to gain greater influential factors on principal component analysis.

Consequently, this paper provides useful information about relationship, but has limit of regional boundary for effectiveness.

Keywords : Sensitiveness, Relationship, Correlation coefficient, Statistical data, Principal component analysis