

단계별 순서를 응용한 첫 일자리에서의 조기퇴직에 대한 영향력 분석 -2009년 대졸자 이동경로조사로부터[†]

정우호¹ · 이성임²

^{1,2}단국대학교 정보통계학과

접수 2010년 9월 26일, 수정 2010년 11월 11일, 게재확정 2010년 11월 23일

요약

본 연구에서는 2007년 한국고용정보원의 설문조사에 의한 <대졸자 이동경로조사 데이터>를 사용하여 첫 일자리에서의 조기퇴직에 대한 영향력을 분석하였다. 조사내용에 의하면 조기퇴직과 관련 있는 설문문항의 수가 매우 크므로, 그 중에서 조기퇴직과 유의한 관련이 있는 문항 즉 설명변수들을 선택하는 데에는 현실적으로 많은 어려움이 존재한다. 본 논문에서는 설명변수의 수가 클 때 자료 분석에서 가능한 모형 선택의 기준을 고찰하고, Shtatland 등 (2003)에서 제안한 모형 선택의 절차를 응용하여 첫 일자리에서의 조기퇴직에 대한 영향력을 분석하였다.

주요용어: 베이지안 정보기준, 변수선택, 아케이케 정보기준, 정보기준.

1. 서론

통계청에서 매년 발표되고 있는 실업률은 2008년 국제 금융위기 이후 상승중이며, 특히 최근 통계청에서 발표한 5년 전국 평균 청년실업률 (15~29세)은 7.7%로 어느 지역의 경우 청년실업률이 9.3%까지 높게 나타나게 됨에 따라, 최근에는 청년들의 노동시장으로의 진입이 어려워지고 있는 실정이다. 취업 시기의 청년들은 자신의 적성과 흥미에 맞는 직장을 선택하기보다는 취업 그 자체에 우선순위를 두고 있다. 또한, 청년 실업률의 증가는 대학졸업을 앞에 둔 청년들로 하여금 취업을 위해 어학점수나 자격증과 같은 스펙 쌓기에 몰두하는 경향을 가져왔다. 따라서 최근의 연구는 실제 취업률에 미치는 주요 요인에 관한 연구가 대부분이었다. 박천수 (2009)는 어학연수 경험의 경우 취업에서 긍정적인 효과를 나타낸다고 하였고, 이만기 (2008)는 자격증 보유여부가 취업에 긍정적인 효과가 있으며, 취업의 질적인 측면에서도 영향을 줄 수 있다고 보고하였다.

한편, 최근 취업난이라고 함에도 불구하고 오히려 청년들이 자신의 직장에 대해 평생직장으로 여기는 비율이 작아지면서 첫 일자리의 근속기간이 짧아지는 경향이 보고되고 있다. 통계청의 2009년 경제활동인구조사 부가조사에 따르면 청년층 (15~29세) 419만 4천명을 대상으로 첫 일자리의 근속기간을 조사하였는데, 근속기간이 1년 미만인 경우가 전체의 46.4%를 차지하였다. 이에 취업에 미치는 주요요인에 관한 연구 뿐 아니라 조기 퇴직에 관련한 주된 요인에 관한 연구도 그 필요성이 대두되고 있다. 기존 연구의 한 예로, 엄동욱 (2008)은 퇴직한 경험이 있는 근로자를 연구대상자로 하여 기업규모에 따라 근

[†] 이 논문은 2009년도 단국대학교 교내연구비에 의하여 연구되었음.

¹ (448-701) 경기도 용인시 수지구 죽전동 126번지, 단국대학교 정보통계학과, 대학원생.

² 교신저자: (448-701) 경기도 용인시 수지구 죽전동 126번지, 단국대학교 정보통계학과, 부교수.

E-mail: silee@dankook.ac.kr

속기간에 차이가 있음을 보고하였다. 본 연구에서는 최근 청년실업률이 높은 이 시기에 청년들이 첫 일 자리에서의 근속기간도 길지 않다는 사실에 주목하여, 2009년 한국고용정보원에서 설문조사한 <대졸자 이동경로조사 데이터>를 바탕으로 청년들의 조기퇴직 여부에 영향을 주는 특성에 대해 비교 분석하고자 한다.

그런데 설문조사 내용에 살펴보면 조기퇴직과 관련하여 이에 영향을 줄 수 있는 많은 설문문항이 있음을 알 수 있고 이것은 설명변수의 수가 많다는 것을 의미한다. 따라서 많은 변수들로부터 가장 영향력 있는 설명변수들의 부분집합을 선택하는 것은 매우 중요한 문제이나 이 문제에 대한 제 1 법칙은 존재하지 아니함이 알려져 있다. 가장 자동화된 모형선택의 방법에는 전진선택법 (forward selection method), 후진소거법 (backward elimination method), 그리고 단계적 선택방법 (stepwise selection method)이 있는데 고민정과 한준태 (2010), 조대현 등 (2009)과 지혜영과 조완형 (2009)도 이러한 단계선택을 사용하였다. 이들 절차에서 유의수준으로 사용될 α 를 선택하는 것은 매우 중요한 문제로 Shtatland 등 (2001), Lee와 Koval (1997), Hosmer와 Lemeshow (1989) 그리고 Steyerberg 등 (2000)등을 참조하면 $0.05 \leq \alpha \leq 0.50$ 임을 알 수 있으나, 이들 중 어떠한 α 에 대해 선택된 모형이라 하더라도 그것이 이론적으로 의미가 있는 것은 아니다 (Shtatland 등, 2003).

이러한 자동화된 순차적 절차와는 다르게 정보기준 (information criterion)을 이용한 모형선택 방법들 또한 널리 사용되고 있다. 이는 가능한 p 개의 설명변수들로부터 고려할 수 있는 모든 가능한 부분집합의 모형에 대해 Akaike (1973, 1974)의 Akaike information criterion (AIC)나 Shwarz (1978)의 Bayesian Information Criterion (BIC)과 같은 정보기준 (information criteria)을 최소로 하는 모형을 선택하는 것으로 제안된다. 그 기준은 다음과 같이 정의된다.

$$AIC = -2\log L_n(p+1) + 2(p+1)$$

$$BIC = -2\log L_n(p+1) + (p+1)\log(n)$$

이 때, $L_n(p+1)$ 는 설명변수 p 개를 갖는 선형모형의 최대우도 함수값을 나타낸다. 이러한 정보기준을 이용한 모형선택에는 두 가지의 중요한 문제가 발생한다. 첫 번째 문제는 Agresti (1996)에서 지적했듯이 변수의 수가 늘어나면 변수들 간의 잠재적인 연관성이나 교호작용이 크게 늘어나 모형의 선택과정이 쉽지 않다는 것이다. 실제로 주효과만을 가정한 모형에서도 가능한 설명변수의 수가 p 일 때 이들 설명변수의 부분집합으로 얻어지는 모든 가능한 모형은 2^p 개가 된다. 예를 들어, 설명변수가 10개 일 때, 모든 비교가능한 모형의 수는 $2^{10} = 1024$ 개가 되고, 본 논문에서 첫 일자리의 조기퇴직과 관련하여 첫 일자리에서의 근무환경이나 근무조건으로 조사된 27개의 설명변수를 고려하면, 비교해야 할 모형의 수는 $2^{27} = 134,217,728$ 개가 되어 이들 모형 전체에서 정보기준 척도를 적용한다는 것은 현실적이지 않은 측면이 있다. 이런 문제는 조장식 (2010)의 분석에서도 나타남을 알 수 있다. 두 번째 문제는 앞의 단계적 모형선택 방법과 마찬가지로 몇몇 특별한 모형의 경우를 제외하고는 모형선택의 일치성과 같은 이론적 정당성을 부여하기 어렵다는 점이다 (Konishi와 Kitagawa, 2008).

위의 두 가지 문제에 대한 해결책으로 본 연구에서는 다음의 두 단계 절차를 고려한다. 첫 단계로 단계적 변수선택 (stepwise variable selection) 방법에서 구할 수 있는 단계별 순서 (stepwise sequence)를 사용하여, 설명변수의 수가 p 일 때 비교해야 할 전체 모형의 수 2^p 개를 줄이는 것이다. 두 번째 단계에서는 정보기준척도를 사용하여 이를 기준으로 최적의 모형을 선택하도록 하는 것이다. 위의 절차를 적용함으로써 일차적으로는 비교하고자 하는 모형의 수를 현실적으로 줄이는 것이 가능하여 계산상 용이하다는 장점이 있다. 다음으로는 정보기준에 의한 방법이 내포모형 (nested model)들 사이의 모형선택에 있어 다양한 모형에 대하여 모형선택의 일치성 (selection consistency)을 지니고 있음이 알려져 있어, 참 모형 (true model)이 우리가 단계적 모형선택에 의하여 사전 선택한 모형들 중에 존재한다면 제안된 절차는 참 모형을 선택함에 있어서 일치성을 지니게 됨을 알 수 있다 (Konishi와 Kitagawa, 2008). 이

에 본 논문에서는 위의 두 가지 단계를 적용하여 최종모형을 선택, 이를 통해서 그 연관성을 분석하고 하고자 한다.

본 논문의 구성은 다음과 같다. 2절에서는 본 논문의 분석자료인 한국고용정보원의 <대졸자 이동경로조사 데이터>에 대해 설명하고 3절에서는 분석절차와 분석결과에 대해 소개하기로 한다. 그리고 마지막 절에서는 결론 및 토의를 하도록 한다.

2. 대졸자 직업이동경로조사 자료

본 논문에서 분석할 자료는 한국고용정보원에서 2007년 설문조사한 ‘대졸자 직업이동경로조사 (Graduates Occupational Mobility Survey: GOMS)’ 자료이다. 이들 자료에서 2006년 8월 졸업자와 2007년 2월 졸업자 및 대학을 졸업한 청년 (16세~29세)들 중 정규직, 상용직으로 첫 일자리의 경험이 있는 7,514명을 선정하였다. 단, 대졸자 이동경로조사의 조사 시점인 2008년 9월까지 첫 직장 근속기간이 계속 되었다면 그 자료는 중도 절단으로 판단하여 근속기간을 설정하였으며, 첫 일자리의 근속기간이 6개월 미만인 경우를 조기퇴직으로 정의하였다. 본 자료 분석에서 관심 있는 첫 일자리 근속기간이 6개월 미만인 연구대상자는 전체의 6.1% (459명)로 나타났다.

표 2.1 첫 일자리에서의 근무환경 및 근무조건

근무환경	만족도	사회보험	혜택
급여	임금 또는 소득	국민연금/특수직역연금	법정퇴직금/퇴직연금
자신의 교육수준과 비교	고용의 안정성	건강보험	유급휴가
희망급여와 비교	직무내용	고용보험	유급출산휴가
신입사원 교육여부	근무환경	산재보험	유급병가
종사자수	근로시간		가족간호휴가
	개인의 발전가능성		시간외수당
	인간관계		상여금
	인사체계		휴업보상
	직장에 대한 사회적 평판		
	일의 자율성과 권한		

GOMS 설문지 문항을 토대로 첫 취업 이후의 근속기간과 관련 가능한 다양한 특성들에 대하여 표 2.1에서 나타난 것처럼 크게 4가지로 분류하였다. 첫 일자리의 근무환경에 대해서는 급여 (단위: 만원), 직장의 업무내용과 자신의 교육 수준 비교, 희망 급여와의 급여수준비교, 신입사원 교육여부, 종사자수 등이 있다. 직장 내 만족도에 관한 조사를 하기 위해, 소득, 고용의 안정성, 그리고 직무내용 등을 포함하여 10가지 항목에 대해 각각의 만족도를 조사하였다. 또한, 직장에서 사회보험 제공여부를 파악하기 위해 4가지 항목을, 휴가나 수당 등에 관련된 직장 내 혜택 유무를 파악하기 위해 8가지 항목에 대해 조사하였다. 다음 절에서는 표 2.1에서 제시한 27개의 변수들이 조기퇴직과 어떠한 관련이 있는지 평가해보기로 한다.

3. 자료분석

3.1. 모형선택절차

서론에서 언급한것처럼 표 2.1에서와 같이 조기퇴직과 관련있는 설명변수가 27개인 경우 최종모형을 선택하는 것은 실제로 매우 어려운 일이다. 서론에서 말한바와 같이주효과만 고려한다고 해도 고려해야 할 모형이 134,217,728개가 됨을 알 수 있다. 이에 본 논문에서는 이인자 교호작용이상은 고려하지 않고 주효과 모형만을 가정하고 분석하기로 한다.

본 자료분석에서와 마찬가지로 설명변수의 p 가 큰 경우에는 이들 설명변수로 가능한 부분집합의 수가 2^p 가 되어 비교해야할 모형의 수가 기하급수적으로 늘어날 뿐 아니라, 이들 부분집합으로 이루어진 모형들이 내포모형 (nested model)이 아니므로 AIC와 같은 정보기준으로 모형을 선택하는 것은 특별한 의미가 없게 된다. 또한 엄청난 수의 모형을 비교하는 것 자체가 매우 소모적인 일이 된다. 이에 Shtatland 등 (2003)이 제안한 것처럼 유의확률을 1에 가깝게 하여 단계별 모형 선택 (stepwise model selection)을 함으로써 비교할 모형의 수를 2^p 개에서 p 개로 하고, 그 후에 p 개의 모형들에 대해 AIC과 BIC를 사용하여 최종 모형을 선택하도록 하는 절차를 소개하기로 한다.

최종모형 선택 절차

1단계: 설명변수 p 개로부터 단계별 선택 방법 (stepwise selection method)을 실시한다. 이 때 변수를 추가선택 (forward selection)하거나 변수를 제거 (backward elimination)할 때 유의확률을 1에 가깝게 하도록 한다. 그 결과로 변수가 가장 유의한 순서대로 선택되는 일련의 단계별 순서 (stepwise sequence)를 얻게 된다.

2단계: 1단계에서 얻은 p 개의 모형에 대해 AIC와 BIC를 구하여 최종 모형을 선택하도록 한다.

3.2. 분석결과

이절에서는 먼저 표 2.1에서 제시한 것처럼 첫 일자리에서의 근무환경 및 근무조건들은 각각 조기퇴직과 관련하여 유의수준 1%에서 조기퇴직과 연관성이 있는 것으로 나타났고 이러한 결과는 표 3.1~표 3.4에 나타나 있다.

먼저 표 3.1에서 첫 일자리 회사의 종사자 수에 따른 조기 퇴직율을 살펴보면 1~29명 규모의 회사에서 조기 퇴직율은 8.8%로 나타났으며 500명이상 규모의 대형회사에서는 1.6%로 그 비율이 낮게 나타났다. 30~99명 규모의 조기 퇴직율은 5.6%로 나타났으며, 100~499명의 조기 퇴직율은 5.1%로 나타나 회사의 종사자수의 규모가 커질수록 조기 퇴직율이 낮아지는 것으로 나타났다. 이러한 결과는 엄동욱 (2008)에서 보고된 중소기업일수록 1년 미만 조기퇴사자의 비율이 높게 나타난다는 연구결과와 일치하는 것으로 나타났다. 첫 직장의 급여에 따른 조기 퇴직율을 살펴보면 전체적으로 급여가 클수록 조기 퇴직율이 낮은 것으로 나타났다. 첫 임금과 희망임금을 비교했을 때에는 임금이 불만족일 경우 조기 퇴직율이 8.5%로 나타났다. 반면에 임금이 만족한 경우 조기 퇴직율이 3.0%로 나타나서 첫 임금이 희망임금과 유사 할수록 조기 퇴직율은 낮은 것으로 나타났다. 첫 직장과 자신의 교육 수준을 비교했을 때 첫 직장의 수준이 낮은 경우에는 조기 퇴직율이 10.2%로 높게 나타났으며, 직장의 수준이 자신의 교육 수준에 견주어 알맞거나 높은 경우에는 4.2%의 조기 퇴직율을 보였다. 또한, 첫 직장에서 신입사원 교육을 하였을 경우 조기 퇴직율이 4.0%인 반면에, 신입사원 조기 교육을 받지 않았을 경우 조기 퇴직율이 10.3%로 높게 나타났다.

표 3.2에서 임금 및 소득에 대해 불만족 한 경우에는 조기 퇴직율은 10.0%로 나타났으며, 보통일 경우 5.6%로 나타났다. 또한 만족할 경우에는 조기 퇴직율이 3.4%로 나타나 임금 및 소득에 만족할수록 조기 퇴직율은 낮게 나타났다. 고용의 안정성에 대해 불만족 할 경우 조기 퇴직율은 15.5%로 나타났으며 보통일 경우 7.6%로 나타났다. 만족할 경우 조기 퇴직율은 3.6%로 나타나서 고용의 안정성에 만족할수록 조기 퇴직율은 낮게 나타났다. 직무내용의 만족도에 대해 불만족 할 경우 조기 퇴직율은 16.6%로 나타났다. 반면 직무내용의 만족도에서 보통으로 응답하였을 경우 조기 퇴직율이 6.5%로 낮게 나타났으며 직무내용에 만족 하였을 경우 조기 퇴직율이 3.2%로 나타나 그 비율이 가장 낮게 나타났

표 3.1 근무환경에 따른 조기 퇴직율 (단위: 명(%))

		조기퇴직	
		예	아니오
종사자 수 규모*	1~29명	274 (8.8)	2835 (91.2)
	30~99명	91 (5.6)	1522 (94.4)
	100~499명	73 (5.1)	1368 (94.9)
	500명 이상	21 (1.6)	1330 (98.4)
급여*	150만원 이하	322 (12.5)	2252 (87.5)
	150만원 초과~250만원 이하	117 (3.2)	3521 (96.8)
	250만원 초과	20 (1.5)	1282 (98.5)
첫 임금과 회망 임금 비교*	불만족	363 (8.5)	3904 (91.5)
	만족	96 (3.0)	3151 (97.0)
첫 직장 과 자신의 교육수준 비교*	낮았다	183 (10.2)	1603 (89.8)
	알맞다	241 (4.9)	4647 (95.1)
첫 직장에서 신입사원 교육 여부*	높다	35 (4.2)	805 (95.8)
	무	260 (10.3)	2265 (89.7)
	유	199 (4.0)	4790 (96.0)
총합		459	7055

주: *는 근속기간이 6개월 미만과 이상의 분포가 유의한 차이를 보인다는 것을 의미함 (p-value<0.01)

다. 직무내용에 만족 할수록 조기 퇴직율이 낮게 나타났다. 근무환경에 대해 불만족 할 경우에는 조기 퇴직율은 12.0%, 보통일 경우 6.8%, 만족하였을 경우 3.5%로 나타나 근무환경에 대해 만족 할수록 조기 퇴직율이 낮게 나타났다. 근로시간에 대하여 불만족한 경우에는 조기 퇴직율이 9.1%, 보통이었을 경우 6.8%, 만족하였을 경우 3.5%로 나타나 다른 만족도 항목과 마찬가지로 만족 할수록 조기 퇴직율이 낮게 나타났다. 하지만 근로시간에 대해 불만족 할 경우의 조기 퇴직율 9.1%는 다른 항목에 대해 불만족 하였을 경우의 조기 퇴직율과 비교 하였을 때 상대적으로 낮게 나타난 것을 확인 할 수 있다. 개인의 발전 가능성에 대해 불만족인 경우에는 조기 퇴직율이 12.3%로 보통이었을 경우 6.8%, 만족하였을 경우 2.8%와 비교하여 조기 퇴직율이 높게 나타났다. 회사 내에서의 인간관계에 대해 불만족인 경우에는 조기 퇴직율은 16.0%, 보통이었을 경우 8.4%, 만족하였을 경우 3.5%로 나타나 인간관계에 대해 만족 할수록 조기 퇴직율이 낮게 나타났다. 첫 직장의 인사체계에 대해 불만족 할 경우에는 조기 퇴직율은 10.9%, 보통이었을 경우 5.8%, 만족 할 경우 3.2%로 인사체계에 만족할수록 조기 퇴직율이 낮게 나타났다. 직장에 대한 사회적 평판에 불만족 할 경우 조기 퇴직율은 13.3%로 나타났다. 또한 보통일 경우 조기 퇴직율은 7.2%, 만족 하였을 경우 조기 퇴직율은 3.4%로 나타나 직장에 대한 사회적 평판에 만족 할수록 조기 퇴직율은 낮게 나타났다. 회사 근무에 대한 자율성에서도 불만족일 경우 조기 퇴직율이 14.9%였다. 반면 자율성에 대해서 연구대상자의 만족도가 보통일 경우 조기 퇴직율은 7.3%였으며 만족 할 경우 2.7%로 그 비율이 가장 낮게 나타났다. 회사 근무에 대한 자율성의 만족도가 높을수록 조기 퇴직율이 낮게 나타났다.

표 3.2 만족도에 따른 조기 퇴직율 (단위: 명(%))

		조기퇴직	
		예	아니오
임금 및 소득*	불만족	216 (10.0)	1951 (90.0)
	보통	157 (5.6)	2637 (94.4)
	만족	86 (3.4)	2467 (96.6)
고용의 안정성*	불만족	134 (15.5)	733 (84.5)
	보통	166 (7.6)	2015 (92.4)
	만족	159 (3.6)	4307 (96.4)

표 3.2 (계속) 만족도에 따른 조기 퇴직율 (단위: 명(%))

		조기 퇴직율 (단위: 명(%))	
		불만족	만족
직무내용*	불만족	168 (16.6)	847 (83.4)
	보통	163 (6.5)	2331 (93.5)
	만족	128 (3.2)	3877 (96.8)
근무환경*	불만족	149 (12.0)	1097 (88.0)
	보통	161 (6.8)	2223 (93.2)
	만족	149 (3.8)	3735 (96.2)
근로시간*	불만족	204 (9.1)	2037 (90.9)
	보통	149 (6.8)	2053 (93.2)
	만족	106 (3.5)	2965 (96.5)
개인의 발전 가능성*	불만족	211 (12.3)	1500 (87.7)
	보통	160 (6.1)	2470 (93.9)
	만족	88 (2.8)	3085 (97.2)
인간관계*	불만족	116 (16.0)	610 (84.0)
	보통	182 (8.4)	1989 (91.6)
	만족	161 (3.5)	4456 (96.5)
인사체계*	불만족	178 (10.9)	1453 (89.1)
	보통	205 (5.8)	3310 (94.2)
	만족	76 (3.2)	2292 (96.8)
직장에 대한 사회적 평판*	불만족	122 (13.3)	794 (86.7)
	보통	212 (7.2)	2715 (92.8)
	만족	125 (3.4)	3546 (96.6)
일의 자율성에 대한 권한*	불만족	164 (14.9)	933 (85.1)
	보통	193 (7.3)	2459 (92.7)
	만족	102 (2.7)	3663 (97.3)
총합		459	7055

주: *는 근속기간이 6개월 미만과 이상의 분포가 유의한 차이를 보인다는 것을 의미함 (p -value<0.01)

표 3.3에서 직장 내 사회보험 가입여부에 따른 조기퇴직비율을 살펴보면, 직장 내 국민연금 및 특수직역연금에 미가입 되었을 때에는 조기 퇴직율이 17.0%이고, 건강보험에 미가입 되었을 경우에는 조기 퇴직율이 16.8%로 높게 나타났다. 또한 고용보험 및 산재보험에 미가입 되었을 때에는 조기 퇴직율이 각각 15.8%와 15.5%로 나타났다. 반면에 4가지 사회 보험에 가입 되었을 때에는 조기 퇴직율은 5.3%로 비교적 낮은 수치를 나타냈다.

표 3.3 사회보험에 따른 조기 퇴직율 (단위: 명(%))

		조기퇴직	
		예	아니오
직장 내 국민연금	미가입	88 (17.0)	429 (83.0)
	가입	371 (5.3)	6626 (94.7)
/특수직역연금*	미가입	86 (16.8)	427 (83.2)
	가입	373 (5.3)	6628 (94.7)
건강보험*	미가입	87 (15.8)	463 (84.2)
	가입	372 (5.3)	6592 (94.7)
고용보험*	미가입	87 (15.5)	473 (84.5)
	가입	372 (5.3)	6582 (94.7)
산재보험*	미가입	87 (15.5)	473 (84.5)
	가입	372 (5.3)	6582 (94.7)
총합		459	7055

주: *는 근속기간이 6개월 미만과 이상의 분포가 유의한 차이를 보인다는 것을 의미함 (p -value<0.01)

표 3.4에서 직장 내 여러 혜택들의 수혜여부에 따른 조기퇴직비율을 살펴보면, 전체적으로 퇴직금, 여러 휴가, 시간외 수당, 상여금 및 휴업보상의 혜택이 제공되는 경우에는 조기 퇴직율이 3%대인 반면에, 이들 혜택이 제공되지 않거나 제공여부에 대한 사실을 모르고 있는 경우에는 조기 퇴직율이 7~15%로

높게 나타난 것을 알 수 있다.

표 3.4 혜택에 따른 조기 퇴직율 (단위: 명(%))

		조기퇴직	
		예	아니오
퇴직금*	미제공	218 (15.6)	1182 (84.4)
	제공	241 (3.9)	5873 (96.1)
유급휴가*	미제공	259 (11.8)	1937 (88.2)
	제공	200 (3.8)	5118 (96.2)
유급출산휴가*	미제공	380 (8.3)	4201 (91.7)
	제공	79 (2.7)	2854 (97.3)
유급병가*	미제공	318 (10.1)	2828 (89.9)
	제공	141 (3.2)	4227 (96.8)
가족간호휴가*	미제공	379 (7.4)	4732 (92.6)
	제공	80 (3.3)	2323 (96.7)
시간외수당*	미제공	305 (8.7)	3203 (91.3)
	제공	154 (3.8)	3852 (96.2)
상여금*	미제공	243 (12.6)	1685 (87.4)
	제공	216 (3.9)	5370 (96.1)
휴업보상*	미제공	382 (7.2)	4890 (92.8)
	제공	77 (3.4)	2165 (96.6)
총합		459	7055

주: *는 근무기간이 6개월 미만과 이상의 분포가 유의한 차이를 보인다는 것을 의미함 (p-value<0.01)

표 3.1~표 3.4에서 제시된 것처럼 주변 근무환경, 만족도, 사회보험, 직장 내 혜택 등 모든 변수가 조기 퇴직과 매우 유의한 주변 연관성 (marginal association)이 있는 것으로 나타났다. 지금부터는 다른 설명변수의 영향력을 제어하고 조기퇴직을 설명하는데 있어 가장 유의한 모형을 통하여 그 관계를 알아 보기로 한다. 이를 위해 로지스틱 회귀분석 (logistic regression)을 실시, 어떤 특성들이 조기 퇴직에 강한 연관성을 있는지 분석하였다.

서론에서 언급한 것처럼 표 2.1에 소개된 27개의 변수들로부터 가능한 모형의 개수는 2^{27} 인 134,217, 728개이다. 이 중 최적화된 모델을 찾는다는 것은 계산절차상 매우 비효율적일 수 있다. 본 연구에서는 이러한 문제를 해결하기 위해서 3.1절에서 소개한 모형선택 절차를 사용하였다. 먼저, 27개의 독립 변수에 대해 변수를 추가선택 (forward selection)할 때의 유의확률을 0.99로 하고, 변수를 제거할 때 (backward elimination)의 유의확률을 0.995로 하여 변수들의 유의하게 선택되는 단계적 순서 (stepwise sequence)를 얻었다 (표 3.5 참조). 조기퇴직과 관련하여 가장 유의한 변수는 급여, 직무내용 그리고 퇴직금 제공여부의 순으로 나타났다. 이들 27개의 모형으로부터 AIC와 BIC를 살펴보면 그림 3.1과 같다.

표 3.5 단계별 변수선택을 통한 단계별 순서 (stepwise sequence)

단계별 모형	선택변수
MODEL1	급여
MODEL2	급여, 직무내용
MODEL3	급여, 직무내용, 퇴직금
MODEL4	급여, 직무내용, 퇴직금, 자율성
MODEL5	급여, 직무내용, 퇴직금, 자율성, 인간관계
MODEL6	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가
MODEL7	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과 희망임금 비교
MODEL8	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수
MODEL9	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가

표3.5 (계속) 단계별 변수선택을 통한 단계별 순서 (stepwise sequence)

MODEL10	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당
MODEL11	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계
MODEL12	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금
MODEL13	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성
MODEL14	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성, 가족간호휴가
MODEL15	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성, 가족간호휴가, 근로시간
MODEL16	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성, 가족간호휴가, 근로시간, 임금만족도
MODEL17	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성, 가족간호휴가, 근로시간, 임금만족도, 고용의안전성
MODEL18	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성, 가족간호휴가, 근로시간, 임금만족도, 고용의안전성, 첫일자리에서신입사원교육여부
MODEL19	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성, 가족간호휴가, 근로시간, 임금만족도, 고용의안전성, 첫일자리에서신입사원교육여부, 근무환경
MODEL20	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성, 가족간호휴가, 근로시간, 임금만족도, 고용의안전성, 첫일자리에서신입사원교육여부, 근무환경, 휴업보상
MODEL21	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성, 가족간호휴가, 근로시간, 임금만족도, 고용의안전성, 첫일자리에서신입사원교육여부, 근무환경, 휴업보상, 국민연금 및 특수직역연금 가입
MODEL22	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성, 가족간호휴가, 근로시간, 임금만족도, 고용의안전성, 첫일자리에서신입사원교육여부, 근무환경, 휴업보상, 국민연금 및 특수직역연금 가입, 산재보험
MODEL23	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성, 가족간호휴가, 근로시간, 임금만족도, 고용의안전성, 첫일자리에서신입사원교육여부, 근무환경, 휴업보상, 국민연금 및 특수직역연금 가입, 산재보험, 건강보험
MODEL24	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성, 가족간호휴가, 근로시간, 임금만족도, 고용의안전성, 첫일자리에서신입사원교육여부, 근무환경, 휴업보상, 국민연금 및 특수직역연금 가입, 산재보험, 건강보험, 교육수준비교
MODEL25	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성, 가족간호휴가, 근로시간, 임금만족도, 고용의안전성, 첫일자리에서신입사원교육여부, 근무환경, 휴업보상, 국민연금 및 특수직역연금 가입, 산재보험, 건강보험, 교육수준비교, 고용보험
MODEL26	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성, 가족간호휴가, 근로시간, 임금만족도, 고용의안전성, 첫일자리에서신입사원교육여부, 근무환경, 휴업보상, 국민연금 및 특수직역연금 가입, 산재보험, 건강보험, 교육수준비교, 고용보험, 직장의사회적평판
MODEL27	급여, 직무내용, 퇴직금, 자율성, 인간관계, 유급출산휴가, 첫임금과희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성, 가족간호휴가, 근로시간, 임금만족도, 고용의안전성, 첫일자리에서신입사원교육여부, 근무환경, 휴업보상, 국민연금 및 특수직역연금 가입, 산재보험, 건강보험, 교육수준비교, 고용보험, 직장의 사회적평판, 유급휴가

AIC 기준으로는 MODEL 13이 가장 좋은 모형으로 나타났으며, BIC 기준으로는 MODEL 6이 가장 좋은 모형으로 선택되었다. 이 두 모형에서 선택된 변수들의 차이는 표 3.5 에서 알 수 있듯이 첫임금과 희망임금비교, 종사자수, 유급병가, 시간외수당, 인사체계, 상여금, 개인발전가능성에 관한 것인데, MODEL 13의 적합결과 회귀계수를 살펴보면, 첫임금과 희망임금비교와 종사자수를 제외하고 나머지 변수들은 유의수준 5%에서 통계적으로 유의하지 않음을 알 수 있었다. 본 연구에서는 두 정보기준 중 BIC를 기준으로 하여 MODEL 6을 초기 퇴직 여부에 관한 영향을 분석하는 최종모형으로 선정하여 표 3.6의 결과를 얻었다.

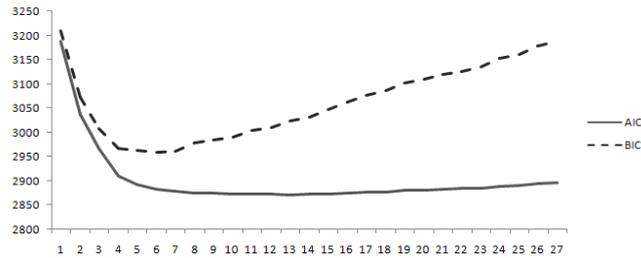


그림 3.1 단계별 AIC 및 BIC 그래프

표 3.6 로지스틱 회귀분석 결과표

변수	수준	$\beta's$	e^β	신뢰구간	유의확률
상수		-4.0747			<.0001
근무 환경	150만원 이하	1.5271	4.605	(2.875, 7.376)	<.0001
	150만원 초과~250만원이하	0.5413	1.718	(1.059, 2.788)	0.0284
	250만원 초과	0			
직무내용	불만족	0.8718	2.391	(1.798,3.180)	<.0001
	보통	0.2130	1.237	(0.953,1.606)	0.1096
	만족	0			
만족도	일의 자율성에 대한 권한	0.9851	2.678	(1.978,3.625)	<.0001
	보통	0.5745	1.776	(1.354,2.330)	<.0001
	만족	0			
인간관계	불만족	0.5958	1.815	(1.353,2.433)	<.0001
	보통	0.4576	1.58	(1.244,2.007)	0.0002
	만족	0			
혜택	미제공 및 모름	0.7392	2.094	(1.676,2.617)	<.0001
	제공	0			
	유급출산휴가	0.4670	1.59	(1.216,2.092)	0.0007
	제공	0			

다른 변수들을 제어하고 해석해 볼 때, 첫 일자리에서의 급여가 150만원 이하인 근로자가 급여가 250만원을 초과하는 근로자에 비해 조기 퇴직을 오즈 (odds)가 약 4.6배 큰 것으로 나타나 역시 급여가 조기퇴직 여부에 커다란 영향이 있다는 것을 알 수 있다. 그 다음 중요한 변수는 직무내용으로 직무내용에 불만족하는 근로자가 만족하는 근로자에 비해 조기 퇴직할 오즈가 약2.4배 큰 것으로 나타나, 직무내용에 불만족할수록 조기 퇴직할 가능성이 큰 것을 뜻한다. 일의 자율성에 대해서도 불만족하는 근로자가 조기 퇴직할 가능성이 큰 것으로 나타났으며, 인간관계 또한 조기퇴직과 관련하여 중요한 요인으로 나타났다. 퇴직금이나 유급출산 휴가를 제공받지 못하는 경우에도 제공받는 경우보다 조기 퇴직할 가능성이 큰 것으로 나타났다.

4. 결론

서론에서 언급하였듯이 구직자들의 노동시장으로의 진입이 어려워짐에 따라 많은 관련 연구자들이 취업에 영향을 미치는 요인들에 대해서 연구가 있어왔고, 청년들의 조기퇴직과 관련한 연구는 미비한 실정이었다. 이에 본 논문에서는 조기 퇴직을 근속기간이 6개월 미만으로 설정한 뒤, 최종 분석한 결과 급여, 직무내용의 만족도, 일의 자율성에 대한 권한, 인간관계, 퇴직금, 유급출산휴가 등이 매우 중요한 요인임을 알아보았다. 분석방법 측면에서는 기존의 박천수 (2009)나 이만기 (2008) 등의 연구에서는 변수

선택을 실시하지 않고, 회귀모형을 적합하여 그 연관성을 평가하였는데, 본 논문에서는 변수선택기법을 적용하였다는 것이 기존 연구와 차이가 있다고 할 수 있다.

본 논문에서는 고려한 설명변수의 수가 27개로 매우 크므로, 주효과만을 대상으로 하여 최종모형 선택절차를 제시하고 모형을 선택하였는데, 앞으로는 2인자 교호작용을 포함한 모형들을 포함하여 최종모형 선택절차를 용이하게 하는 절차 또한 연구할 필요가 있겠다.

참고문헌

- 고민정, 한준태 (2010). 주요 위험요인별 허혈성심질환 사망위험도 분석. <한국데이터정보과학회지>, **21**, 201-209.
- 김정훈 (2010). 대졸자들의 첫 일자리 근속기간에 관한 연구. <한국고용정보원 고용동향조사 (OES, YP, GOMS) 심포지엄>, 344-354.
- 박가열, 임영식 (2007). 학교에서 일터로의 이행 초기 직무만족에 영향을 미치는 요인에 관한 연구. <제6회 산업·직업별 고용구조조사 및 청년패널 심포지엄>, 296-307.
- 박천수 (2009). 대학생의 어학연수가 노동시장 이행에 미치는 영향에 관한연구. <제1회 대졸자 직업이동 경로조사 심포지엄 논문집>, 173-96.
- 엄동욱 (2008). 대졸신규대졸취업자의 직장이동과 노동시장 성과: 기업 HRM에 의한 합의를 중심으로. <제1회 대졸자 직업이동경로조사 심포지엄 논문집>, 한국고용정보원.
- 이만기 (2008). 자격취득자의 실태와 자격의 노동시장 효과에 관한 분석. <제1회 대졸자 직업이동 경로조사 심포지엄 논문집>, 199-221.
- 조대현, 김병수, 석경하, 이종언, 김종성, 김선화 (2009). 화장품구매 자료를 통한 고객 구매형태 분석. <한국데이터정보과학회지>, **20**, 615-627.
- 조장식 (2010). 학업성취도에 대한 대입전형 요인들의 영향력 분석. <한국데이터정보과학회지>, **21**, 729-736.
- 지혜영, 조완형 (2009). 인터넷 쇼핑몰에서 구매품목을 이용한 고객의 예측모델 설계. <한국데이터정보과학회지>, **20**, 25-37.
- 최중후, 한상태, 강현철, 김은석 (1998). <AnswerTree를 이용한 데이터마이닝 의사결정나무분석>, SPSS아카데미, 서울.
- Agresti, A. (1996). *An introduction to categorical data analysis*, John Wiley and Sons, Inc., New York.
- Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. *Second International Symposium on Information Theory* edited by B. N. Petrov and F. Csaki. Budapest: Akademiai Kiado, 267-81.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control AC*, **19**, 716-723.
- Hosmer, D. W. and Lemeshow, S. (1989). *Applied logistic regression*, John Wiley and Sons, Inc., New York.
- Konishi, S. and Kitagawa, G. (2008). *Information criteria and statistical modelling*, Springer-Verlag, New York.
- Lee, K. and Koval, J. J. (1997). Determination of the best significance level in forward logistic regression. *Communications in Statistics - Simulations*, **26**, 559-575.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461-464.
- Shtatland, E. S., Cain E. and Barton, M. B. (2001). The perils of stepwise logistic regression and how to escape them using information criteria and the Output Delivery System. *SUGI'26 Proceedings, Paper 222-26*. Cary, NC: SAS Institute Inc.
- Shtatland, E. S., Kleinman, K. and Cain, E. M. (2003). Stepwise methods in using SAS PROC LOGISTIC and SAS ENTERPRISE MINER for prediction. *SUGI'28 Proceedings, Paper 258-28*. Cary, NC: SAS Institute Inc.
- Steyerberg, E. W., Eijkemans, M. J. C., Harrell Jr, F. E. and Habbema, J. D. F. (2000). Prognostic modeling with logistic regression analysis: A comparison of selection and estimation methods in small data sets. *Statistics in Medicine*, **19**, 1059-1079.

Analysis of the impact on quitting one's first job using the stepwise sequence - based on graduates occupatinal mobility survey

Wooho chung¹ · Sungim Lee²

¹²Department of Statistics, Dankook University

Received 26 September 2010, revised 11 November 2010, accepted 23 November 2010

Abstract

In this paper, we analyze the impact on quitting one's first job based on "Graduates Occupational Mobility Survey" data given by Korea Employment Information Service. According to the survey, there are a large number of questionnaires on quitting one's first job and so it is not easy to choose among them. We will investigate model selection criteria and apply the procedure proposed by Shtatland *et al.* (2003) to identify the final model.

Keywords: AIC, BIC, information criterion, model selection.

¹ Master's course, Department of Information Statistics, Dankook University, 126, Jukjeon-dong, Suji-gu, Yongin-si, Gyeonggido 448-701, Korea.

² Corresponding author: Associate professor, Department of Information Statistics, Dankook University, 126, Jukjeon-dong, Suji-gu, Yongin-si, Gyeonggido 448-701, Korea. E-mail: silee@dankook.ac.kr