

의사결정나무모형을 사용한 성인 생애주기별 취업 영향요인 분석

곽민정¹ · 이성석²

¹평택대학교 데이터정보학과 · ²서원대학교 경영학과

접수 2016년 10월 11일, 수정 2016년 11월 14일, 게재확정 2016년 11월 21일

요약

세계적으로 유가하락과 더불어 경제가 침체되면서 우리나라도 저성장의 기조를 보이고 있고, 노동 시장에서는 취업난이 가중되고 있으므로 취업영향 요인을 파악하여 적절한 취업 정책을 수립하는 것이 절실한 현실이다. 따라서 본 연구에서는 제17차년도 노동패널자료를 사용하여 취업에 영향을 미치는 요인을 파악하고자 한다. 성인 생애주기는 청년기, 중장년기와 노년기로 구분하였으며 취업에 영향을 미치는 요인으로 인구통계학적 변수, 직업관련 변수 그리고 건강관련 변수를 고려하였다. 의사결정나무분석을 사용하여 분석한 결과 청년기에는 학력이 가장 중요한 요인이었으며, 중장년기에는 가장 중요한 요인이 성별이었고, 남성의 경우 건강상태, 여성의 경우, 직업훈련경험, 연령, 건강상태의 순으로 나타났다. 노년기에도 성별이 가장 중요한 요인이었고 그 다음으로 건강상태, 학력 등의 순으로 나타났다.

주요용어: 생애주기, 의사결정나무모형, 패널자료.

1. 머리말

세계적으로 유가하락과 더불어 경제가 침체되면서 우리나라도 저성장의 기조를 나타내고 있다. 최근에 경제활동인구가 증가되면서 취업자도 증가하고 있지만 실업자나 비경제활동인구도 증가하고 있어서 미취업자도 동시에 증가하고 있는 상황이다. 최근에 대학졸업자를 비롯한 청년층의 미취업문제가 부각되고 있을 뿐만 아니라 평균수명의 증가와 더불어 정년이 연장되는 등의 추세와 함께 고령층의 미취업 문제가 심각해지고 있는 상황이다. 취업 문제는 청년층이나 여성 또는 장애인 등과 같은 특수한 집단만의 문제가 아니라 전 세대에 걸쳐 발생하고 있다.

인간의 생애주기 (human life cycle)를 두 가지로 나누어 보면, 미성년기와 성년기로 나누어 볼 수 있다. 취업은 성년기의 중요 활동이며, 성년기를 생물학적인 연령에 따라서 다시 구분하여 보면, 청년기, 장년기 또는 중년기, 노년기로 구분할 수 있다 (Jung, 2012). 생애주기는 생물학에 기초한 개념으로, 사회학, 가족학 등에서 활용되고 있으며, 최근에는 경제학, 경영학, 종교학 등 많은 분야에서 응용되고 있다.

취업여부의 결정요인을 탐색하는 기존의 연구를 살펴보면 청년층 (Lee 등, 2010), 대졸자 (Choi와 Kim, 2013), 중고령자 (Sung과 Ahn, 2006; Yum, 2008) 등과 같이 대부분 특정 연령집단에 대하여 이루어진 경우가 많으며, 여성 (Lee, 2004), 장애인 (Oh, 2008), 저소득층 (Ryu, 2012) 등과 같이 특정한

¹ (17869) 경기도 평택시 서동대로 3825, 평택대학교 데이터정보학과, 부교수.

² 교신저자: (28674) 충북 청주시 서원구 무심서로 377-3, 서원대학교 경영학과, 교수.

E-mail: ssrhee@seowon.ac.kr

속성을 가진 집단으로 제한된 경우가 많았다. 본 연구는 분석대상자들을 성인의 생애주기에 따라 세 단계, 즉 청년기, 중장년기, 노년기로 구분하였고, 성인의 생애주기에 영향을 줄 수 있는 요인으로 인구통계학적 변수와 직업관련 변수를 사용하였으며 건강관련 변수를 추가적인 요인으로 고려하였으며, 의사결정나무모형을 이용한 데이터마이닝 분석기법을 적용하였다.

본 연구의 목적은 기본적으로 취업결정에 영향을 미치는 요인들을 탐색하는 것이며, 동시에 취업의 문제가 성인의 생애주기에 걸친 것으로 인식하고 이에 적합한 요인을 찾아낼 뿐만 아니라 무슨 요인이 중요한지를 의사결정모형을 이용하여 규명하고자 한다.

2. 연구방법

2.1. 연구자료

본 연구에서는 한국노동연구원의 제17차년도 한국노동패널조사의 개인자료를 이용하였다. 이 자료에서 개인자료는 총 13,169명이지만, 먼저 휴학생을 포함한 재학생(고등학생, 대학생, 대학원생)은 미취업의 원인임이 분명하므로 이에 해당하는 1,105명(결측치 2명 포함)을 제외시켰으며, 다음으로 미성년자와 노령자는 취업을 하기에 부적합할 것으로 예상되므로 18세 이하 또는 71세 이상의 1,893명을 추가적으로 제외시켰다. 따라서 본 연구의 분석대상은 재학생을 제외한 19세부터 70세까지의 10,171명이다. 또한 성인의 생애주기를 생물학적 연령에 기초하여 세 개로 나누었고, 19세부터 29세까지를 청년기, 30세부터 59세까지를 중장년기, 60세부터 70세까지를 노년기로 명명하였다. 성인 생애주기에 따라, 취업여부별 분포를 살펴보면 Table 2.1과 같다. 성인 생애주기별 빈도는 청년기 788명(7.7%), 중장년기 7,375명(72.5%), 노년기 2,008명(19.7%)으로 나타났으며, 취업여부별 빈도는 취업자 7,105명(69.9%), 미취업자 3,066명(30.1%)으로 나타났다. 여기서 미취업자가 낮게 나타난 것은 이미 휴학생을 포함한 재학생, 18세 이하, 71세 이상의 제외 조건에 의하여 미취업자 중 상당수가 제외된 결과로 사료된다.

Table 2.1 Adult groups by life cycle

		Adult groups by life cycle			Total
		Young adult	Adult	Old adult	
Employment ($p < 0.0001$)	Employed	526 (7.4%)	5,517 (77.7%)	1,062 (14.9%)	7,105 (100.0%)
	Unemployed	262 (8.5%)	1,858 (60.6%)	946 (30.9%)	3,066 (100.0%)
	Total	788 (7.7%)	7,375 (72.5%)	2,008 (19.7%)	10,171 (100.0%)

2.2. 변수의 설정

목표변수는 취업여부이고, 입력변수는 첫째, 인구통계학적인 변수로서 성별, 연령, 학력, 부 또는 모의 생존여부, 사회경제적 지위, 사회경제적 지위의 상승가능성 등 6개를 사용하였고, 둘째, 직업과 관련된 변수로서 직업훈련여부, 자격증취득유무, 좋은 일자리 기준 등 3개를 사용하였으며, 셋째, 건강과 관련된 변수로서 현재의 건강상태, 1년 전과 비교한 건강상태 등 2개를 사용하였다. 인구통계학적 특성, 직업관련 변수 또는 건강관련 변수와 취업여부의 관련성은 카이제곱검정을 사용하여 분석하였다. 취업여부(목표변수)를 결정하는 영향변수(입력변수)가 무엇인가를 탐색하기 위하여 SAS의 E-miner(Ver. 13.1)를 이용하여 의사결정나무모형 분석을 실시하였으며 목표변수의 결측치도 하나의 범주로 간주하여 분석하였다.

2.3. 의사결정나무모형

의사결정나무는 적절한 의사결정규칙에 따라 나무구조의 모형으로 도식화하여 분류와 예측을 수행하는 방법이다. 이는 연구자가 그 과정을 쉽게 이해하고 설명할 수 있고(Breiman 등, 1993; Choi 등, 2015) 의사결정나무분석에서 목표변수에 영향을 미치는 요인을 찾아내고 각 요인의 중요도를 결정할 수

있으며, 각 과정에서 영향을 미치는 중요한 요인에 따라 발생비율이 어떻게 달라지는 지를 도식화된 흐름도로 표현할 수 있다. 또한, 비모수적 방법으로서 선형성과 같은 가정을 필요로 하지 않는 장점을 가지고 있으며 요인들간 상호작용을 파악하는데 용이하다 (Cho, 2010; Cho, 2014; Jung 등, 2013). 본 연구는 취업에 영향을 미치는 요인을 탐색하는 것이 주된 목적으로 제17차년도 한국노동패널자료를 사용하여 취업여부에 영향을 미치는 요인을 파악하고자 한다. 한국노동패널자료는 설문조사자료이며 대부분 범주형 자료로 구성되어 있어 로지스틱회귀모형을 사용하여 각 요인의 효과크기를 측정하고 해석하는 것은 한계가 있으며 취업자와 미취업자를 판단하고 예측하는 취지보다는 영향요인을 탐색하고 각 요인과의 상호작용을 살펴보고자 의사결정나무모형을 사용하였다.

3. 연구결과

3.1. 생애주기별 인구통계학적 특성

분석대상인 10,171명의 생애주기별 인구통계학적인 분포특성은 Table 3.1에 제시되었다. 청년기의 취업자의 인구통계학적 특성을 살펴보면 성별에서 남성은 45.1%, 여성은 54.9%로 여성의 비율이 약간 높게 나타났다. 학력은 고졸이하가 19.4%, 대졸이상이 80.6%로 대부분 대졸이상의 학력을 보유하고 있음을 알 수 있다. 부 또는 모의 생존은 97.3%이며 사회경제적 지위는 중이 70.4%로 대부분을 차지하였다.

중장년기의 취업자의 인구통계학적 특성은 성별에서 남성이 60.7%, 여성은 39.3%로 다른 생애주기에 비하여 남성의 비율이 더욱 높게 나타났다. 학력은 고졸이하가 52.3%, 대졸이상이 47.7%로 고졸이하와 대졸이상의 학력이 비슷함을 알 수 있다. 부 또는 모의 생존은 79.2%이며 사회경제적 지위는 중이 69.6%로 대부분이었다.

노년기의 취업자 성별은 남성이 59.3%, 여성은 40.7%로 남성의 비율이 약간 높게 나타났다. 학력은 고졸이하가 89.3%로 청년기나 중장년기에 비하여 저학력의 특징을 보이고 있다. 부 또는 모의 생존은 22.4%이며 사회경제적 지위는 중이 60.4%이었다.

Table 3.1 Descriptive statistics for demographic variables according to adult life cycle and employment status

Adult life cycle	Employment	Young adult		Adult		Old adult	
		employed	unemployed	employed	unemployed	employed	unemployed
Gender	male	237 (45.1)	97 (37.0)	3351 (60.7)	268 (14.4)	630 (59.3)	293 (31.0)
	female	289 (54.9)	165 (63.0)	2166 (39.3)	1590 (85.6)	432 (40.7)	653 (69.0)
	$\chi^2 = 4.62, p = 0.0316$		$\chi^2 = 119.02, p < 0.0001$		$\chi^2 = 161.90, p < 0.0001$		
Education	less	1 (0.2)	5 (1.9)	698 (12.7)	299 (16.1)	648 (61.0)	582 (61.5)
	high school	101 (19.2)	93 (35.5)	2185 (39.6)	783 (42.2)	300 (28.3)	259 (27.4)
	higher	424 (80.6)	164 (62.6)	2633 (47.7)	775 (41.7)	114 (10.7)	105 (11.1)
	$\chi^2 = 33.25, p < 0.0001$		$\chi^2 = 25.28, p < 0.0001$		$\chi^2 = 0.22, p = 0.8967$		
parents	yes	512 (97.3)	258 (98.5)	4369 (79.2)	1463 (78.7)	238 (22.4)	183 (19.3)
	no	14 (2.67)	4 (1.5)	1145 (20.8)	395 (21.3)	824 (77.6)	763 (80.7)
	$\chi^2 = 1.01, p = 0.3151$		$\chi^2 = 0.21, p = 0.6505$		$\chi^2 = 2.84, p = 0.0920$		
Socio-economic status	high	7 (1.3)	4 (1.6)	77 (1.4)	29 (1.6)	15 (1.4)	18 (1.9)
	middle	369 (70.4)	168 (65.4)	3819 (69.6)	1201 (64.8)	638 (60.4)	554 (58.6)
	low	148 (28.2)	85 (33.1)	1595 (29.0)	624 (33.7)	403 (38.2)	374 (39.5)
$\chi^2 = 2.05, p = 0.3593$		$\chi^2 = 14.62, p = 0.0007$		$\chi^2 = 1.23, p = 0.5395$			

직업관련 변수들을 살펴보면 Table 3.2와 같다. 청년기의 취업자의 대부분 (92.9%)은 직업훈련 경험이 전무하며 자격증 취득비율도 매우 낮은 편이었다. 좋은 일자리 기준은 안정적인 일자리를 40.5%로 가장 선호하였고 그 다음으로 자신의 적성과 취향에 맞는 일자리, 장래의 발전 가능성이 높은 일자리, 사회적으로 인정받는 일자리의 순이었다.

중장년기의 경우에도 청년기와 마찬가지로 직업훈련경험이나 자격증 취득이 매우 낮은 경향을 보였다. 좋은 일자리 기준은 52.6%로 안정적인 일자리를 가장 선호하였고 그 다음으로 자신의 적성이나 취향에 맞는 일자리, 임금 또는 수입이 많은 일자리 순으로 나타났다.

노년기에도 청년기나 중장년기와 같이 직업훈련경험이나 자격증 취득이 매우 낮은 경향을 보였다. 좋은 일자리 기준은 중장년기와 마찬가지로 안정적인 일자리, 자신의 적성이나 취향에 맞는 일자리, 임금 또는 수입이 많은 일자리 순으로 선호하였다. 전 세대에서 모두 일자리의 안정성을 가장 중요시하는 것을 알 수 있다.

Table 3.2 Descriptive statistics for occupational variables according to adult life cycle and employment status

Adult life cycle		Young adult		Adult		Old adult	
Employment		employed	unemployed	employed	unemployed	employed	unemployed
Job training	before	35	4	367	14	35	0
		(6.7)	(1.5)	(6.7)	(0.8)	(3.3)	(0.0)
	present	2	2	20	16	2	1
		(0.4)	(0.8)	(0.4)	(0.9)	(0.2)	(0.1)
never	488	256	5118	1828	1024	945	
		(92.9)	(97.7)	(93.0)	(98.4)	(96.5)	(99.9)
		$\chi^2=10.24, p=0.0060$		$\chi^2=105.23, p<0.0001$		$\chi^2=32.02, p<0.0001$	
Certificate	yes	14	3	48	16	1	3
		(2.7)	(1.2)	(0.9)	(0.9)	(0.1)	(0.3)
	no	512	259	5461	1842	1060	943
		(97.3)	(98.9)	(99.1)	(99.1)	(99.9)	(99.7)
		$\chi^2=1.91, p=0.1675$		$\chi^2=0.00, p=0.9674$		$\chi^2=1.25, p=0.2638$	
Criteria for good job	stability	212	111	2885	891	633	549
		(40.5)	(42.7)	(52.6)	(48.1)	(59.9)	(58.0)
	social reputation	54	24	452	181	84	86
		(10.3)	(9.3)	(8.2)	(9.8)	(8.0)	(9.1)
	future development	75	43	524	222	77	98
		(14.3)	(16.6)	(9.6)	(12.0)	(7.3)	(10.4)
aptitude	119	52	886	293	126	122	
	(22.7)	(20.1)	(16.2)	(15.8)	(11.9)	(12.9)	
income	53	22	657	227	110	83	
	(10.1)	(8.5)	(12.0)	(12.3)	(10.4)	(8.8)	
autonomy	11	7	80	39	26	8	
	(2.1)	(2.7)	(1.5)	(2.1)	(2.5)	(0.9)	
		$\chi^2=2.33, p=0.8015$		$\chi^2=21.01, p=0.0008$		$\chi^2=15.89, p=0.0072$	

건강상태에 대하여 살펴보면 Table 3.3과 같다. 청년기에 취업자는 대부분 건강한 편이라고 답하였고 1년 전과 비교한 건강상태도 대부분 비슷하거나 약간 나아졌다고 응답하였으며 중장년기의 건강상태에 대한 응답은 청년기와 비슷한 경향을 보이고 있다. 노년기의 취업자중에서 건강하지 않은 편이라고 한 응답자가 13.5%이었으며 1년 전과 비교한 건강상태에서 나빠진 편이라는 응답자가 10.5%인 특징을 보이고 있다.

Table 3.3 Descriptive statistics for health status variables according to adult life cycle and employment status

Adult life cycle		Young adult		Adult		Old adult	
Employment		employed	unemployed	employed	unemployed	employed	unemployed
Health status	excellent, very good	371 (70.5)	195 (74.4)	3460 (62.8)	1019 (54.8)	448 (42.22)	275 (29.1)
	good	151 (28.7)	55 (21.0)	1855 (33.7)	563 (30.0)	470 (44.3)	370 (39.1)
	fair, poor	4 (0.8)	12 (4.6)	195 (3.5)	276 (14.9)	143 (13.5)	301 (31.8)
		$\chi^2 = 16.92, p = 0.0002$		$\chi^2 = 297.56, p < 0.0001$		$\chi^2 = 103.28, p < 0.0001$	
Health status compared to 1-year before	better	42 (8.0)	10 (3.82)	261 (4.7)	61 (3.3)	35 (3.3)	16 (1.7)
	the same	480 (91.3)	244 (93.1)	5077 (92.1)	1620 (87.2)	915 (86.2)	708 (74.8)
	worse	4 (0.8)	8 (3.05)	173 (3.2)	177 (9.5)	111 (10.5)	222 (23.5)
		$\chi^2 = 10.71, p = 0.0047$		$\chi^2 = 129.79, p < 0.0001$		$\chi^2 = 64.10, p < 0.0001$	

3.2. 청년기의 취업결정요인에 대한 의사결정나무 분석결과

Figure 3.1에 나타나 있는 의사결정나무 분석결과를 살펴보면, 취업여부에 가장 큰 영향을 미치는 요인은 학력이었다. 구체적으로 살펴보면, 청년기에 해당하는 분석대상은 788명이며 이중 취업자 비율은 66.8%를 차지했다. 다음으로 학력이 고졸이하의 경우는 200명으로 이중 취업자 비율은 51.0%인 반면, 대졸 (전문대학 포함)이상의 경우는 588명이고 이중 취업자 비율은 72.1%로 높게 나타났다. 이러한 결과는 정보통신기술의 혁명과 지식산업의 고도화로 인하여 4차 산업혁명의 시대를 맞이하고 있는 요즘은 낮은 교육수준의 인력보다는 높은 교육수준의 인력에 대한 산업계의 수요가 많아지는 추세에 따른 것으로 볼 수 있다.

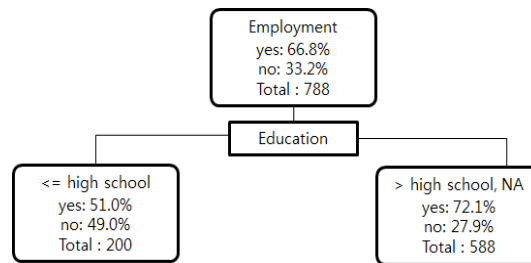


Figure 3.1 Decision tree result for young adult (misclassification rate=33.2%)

3.3. 중장년기의 취업결정요인에 대한 의사결정나무 분석결과

Figure 3.2에 나타나 있는 의사결정나무 분석결과를 살펴보면, 취업여부에 가장 큰 영향을 미치는 요인은 성별이었다. 남성의 경우 건강상태만이 유의한 요인으로 선택되었으나 여성의 경우 직업훈련경험의 유무, 연령 그리고 건강상태의 순으로 유의하게 선택되었다.

구체적으로 살펴보면, 중장년기에 해당하는 분석대상은 7,375명이며 이중 취업자 비율은 74.8%로 나타났다. 남성의 경우 취업자 비율은 92.6%로 상당히 높게 나타났으나 건강상태가 좋지 않은 경우 46.8%로 낮아졌다. 여성의 경우 취업자 비율은 57.7%이며, 직업훈련경험이 있는 경우 92.4%로 매우 높게 나타났다. 직업훈련을 현재 받고 있거나 경험이 없는 여성의 경우 연령이 43세보다 크고, 건강상태가 좋은 경우 취업자 비율이 64.3%로 높은 경향을 보이고 있다.

이러한 결과는 중장년기에는 남성과 여성 모두 건강상태가 취업여부를 결정짓는 데 중요한 요인으로 작용하고 있으며 특히 여성의 경우에는 직업훈련을 강화하는 정책이 더욱 효과적임을 시사하고 있다.

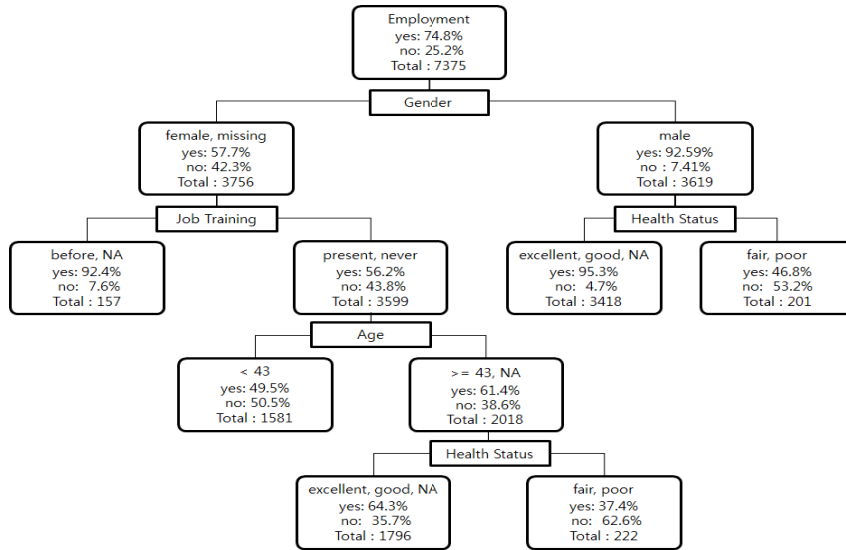


Figure 3.2 Decision tree result for adult (misclassification rate=23.5%)

3.4. 노년기의 취업결정요인에 대한 의사결정나무 분석결과

Figure 3.3의 의사결정나무 분석결과에서, 취업여부에 가장 큰 영향을 미치는 요인은 성별이었으며 그 다음으로 건강상태, 학력 등이었다.

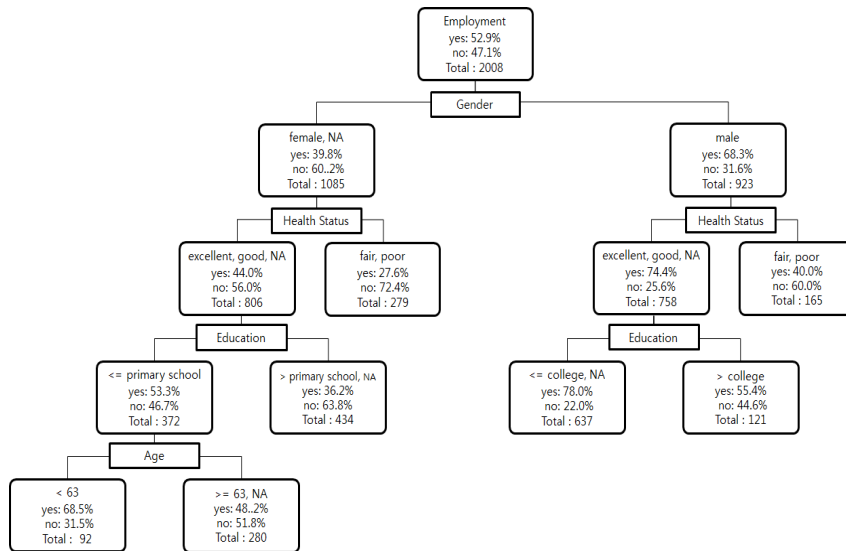


Figure 3.3 Decision tree result for old adult (misclassification rate=32.7%)

구체적으로 살펴보면, 노년기에 해당하는 분석대상은 2,008명이며 이중 취업자 비율은 52.9%로 나타났다. 남성의 취업자 비율은 68.3%인 반면 여성은 39.8%로 상반되게 나타났다. 남성의 건강상태가 좋은 경우 취업자 비율은 74.4%이었고 학력이 전문대졸이하의 경우 취업자 비율은 78.0%로 나타났다. 여성의 건강상태가 좋은 경우 취업자 비율은 44.0%로 약간 높게 나타났으며 특히, 학력이 초등학교졸업이 하이고 연령이 62세 이하인 경우에 68.5%로 상당히 높게 나타났다. 이러한 결과를 요약해보면 노년기에는 남성의 경우 건강상태가 좋고, 학력이 낮을수록, 그리고 연령이 낮을수록 취업자 비율이 높게 됨을 알 수 있다. 특히, 노년기에는 취업에 있어서 건강관리가 매우 중요하다는 것을 보여준다.

4. 결론

본 연구의 목적은 제17차년도 한국노동패널자료를 이용하여 취업결정에 영향을 미치는 요인들을 찾아내는 것으로 인구통계학적인 요인, 직업관련 요인, 건강관련 요인 등을 탐색하였다. 또한 성인의 생애주기를 청년기, 중장년기, 노년기로 구분하고 연령에 따라 연구대상을 세 집단으로 나누었다.

본 연구를 통하여 얻은 결론은 다음과 같다.

첫째, 취업자의 인구통계학적 특성은 남성의 비율이 청년기에 비하여 중장년기와 노년기에 높게 나타났고, 대졸자 비율은 청년기에, 중졸자 비율은 노년기에 높게 나타났다. 취업자 중에서 사회경제적 지위는 대부분 중산층으로 나타났다.

둘째, 취업자 중에서 직업훈련경험이나 자격증 취득자 비율은 매우 낮았으며, 좋은 일자리의 기준으로 전 세대에서 안정적인 일자리를 가장 선호하는 것으로 나타났다. 청년기에는 자신의 적성과 취향에 맞는 일자리나 장래의 발전가능성이 높은 일자리를 원하는 응답자가 상대적으로 많았다.

셋째, 취업자 중에서 건강상태가 보통이거나 건강한 편이란 응답자가 대부분이며, 노년기에는 건강하지 않은 편이라는 비율이 상대적으로 높게 나타났다.

넷째, 청년기에는 학력이 가장 중요한 요인이며, 대졸이상의 경우에 취업자 비율이 높게 나타났다.

다섯째, 중장년기에는 성별이 가장 중요한 요인으로, 남성이 여성보다 취업자 비율이 높았으며 남성은 건강상태가 좋은 경우에, 여성은 직업훈련경험이 있는 경우에 취업자 비율이 높은 것으로 나타났다. 특히 직업훈련경험이 없는 여성은 연령이 높고 건강상태가 좋을수록 취업자 비율이 높았다.

여섯째, 노년기에도 성별이 가장 중요한 요인이며, 남성이 여성보다 취업자 비율이 높았다. 남성이나 여성 모두 공통적으로 건강상태가 좋고 학력이 낮을수록 취업률이 높은 것으로 나타났다. 이는 학력이 낮은 경우 공평한 생활 때문에 취업을 유지하고 있는 것으로 해석할 수 있다.

본 연구에서 성인의 생애주기를 청년기, 중장년기, 노년기로 구분하고, 생애주기별로 취업결정요인들을 의사결정나무모형으로 분석해본 결과 중요한 취업결정요인들이 생애주기별로 서로 상이할 뿐만 아니라 그 중요성의 순서도 다르게 나타났다. 따라서 취업정책도 생애주기별로 달리 수립되어야 할 것으로 사료된다. 앞으로 보다 많은 취업결정요인을 고려하여 보다 타당성이 높은 모형을 수립하여야 할 것이며, 실업자나 미취업자에 대하여 취업을 용이하게 할 수 있는 구체적인 방안을 모색하여야 할 것이다.

References

- Breiman, L., Friedman, J., Stone, C. and Olshen, R. A. (1993). *Classification and regression trees*, CRC Press, New York.
- Cho, J. S. (2010). The influence analysis of admission variables on academic achievements. *Journal of the Korean Data & Information Science Society*, **21**, 729-736.
- Cho, J. S. (2014). Analysis of employee's characteristic using data visualization. *Journal of the Korean Data & Information Science Society*, **25**, 727-736.

- Choi, K. L. and Kim, B. (2013). A study on the relationship between job transition and personal attributes using multiple logit model. *Journal of the Korean Data Analysis Society*, **15**, 799-811.
- Choi, Y. J., Kwak, M. J. and Yoon, M. (2015). The diffusion and policy options of the diagnostic imaging technologies in Korea. *Journal of the Korean Data & Information Science Society*, **26**, 179-185.
- Jung, I. H. (2012). A study on self-esteem and influencing factors of adults by life cycle: comparison of young, middle-aged and elderly. *Korean Review of Crisis and Emergency Management*, **8**, 235-250.
- Jung, J. H. and Min, D. K. (2013). The study of foreign exchange trading revenue model using decision tree and gradient boosting. *Journal of the Korean Data & Information Science Society*, **24**, 161-170.
- Lee, S. (2004). A study on the cause of unemployment for women: Based on reservation wage and market wage. *Korean Journal of Labour Economics*, **27**, 135-164.
- Lee, Y. S., Dong, S. O., Jung, Y. A., Kang, C. and Kim, K. K. (2010). Background of successful Employment after graduation. *Journal of the Korean Data Analysis Society*, **12**, 1523-1533.
- Oh, S. R. (2008). A study about the predictor variables of employment of persons with disabilities. *Social Welfare Policy*, **34**, 255-275.
- Sung, J. M. and Ahn, J. Y. (2006). What makes the older work for satisfactory? *Journal of Labor Policy*, **6**, 39-74.
- Ryu, G. R. (2012). Analysis of labor market activation policy and its employment outcomes: The effects of employment and tailored social service provision. *Korean Social Policy Review*, **19**, 149-183.
- Yum, D. W. (2008) The determinants of the older's decision to work: Focusing on 1st Korean retirement and income study. *Journal of Labor Policy*, **8**, 17-38.

Finding factors on employment by adult life cycle using decision tree model

Minjung Kwak¹ · Sung-Suk Rhee²

¹Department of Data Information, Pyeongtaek University

²Department of Business Administration, Seowon University

Received 11 October 2016, revised 14 November 2016, accepted 21 November 2016

Abstract

Due to global economy recession with low oil price, Korea has stepped into a stage of sluggish development, and the unemployment has become a major issue. Hence, we study various risk factors influencing on unemployment using the Korean labor and income panel data of 2014. We first divide the adult life cycle into three categories, such as young adult, adult, and old adult. Then we consider demographic variables, occupational variables and health related variables as risk factors. The decision tree models have shown that education and gender are the most important factors respectively in young adult group and in adult group. Gender, health status, and education are influential factors in old adult group.

Keywords: Decision tree, life cycle, panel data.

¹ Associate professor, Department of Data Information, Pyeongtaek University, Pyeongtaek 17869, Korea.

² Corresponding author: Professor, Department of Business Administration, Seowon University, Cheongju 28674, Korea. E-mail: ssrhee@seowon.ac.kr