

대학입시제도와 관련한 諸假說의 검정

— 연세대학교의 입시자료를 중심으로 —

金秉洙 · 金善柱

1994학년도부터 대학입시에서 본고사가 부활되면서 무시험 특별전형제도가 도입되었다. 이와 함께 수능능력시험과 고교내신 성적이 입시의 주요 변수로 작용하였다. 연세대학교는 1994학년도에 입학정원의 20%를 특별전형으로, 1995학년도에는 40%를 특별전형으로 선발한 바 있다. 1994·1995학년도 수험생의 입시자료와 1994학년도 입학생의 1년간 GPA가 알려진 1995년 초에 특별전형의 효과를 분석하면서, 입시와 관련된 다음 문제들을 규명하였다. ① 수능능력시험과 본고사와의 상관관계는 어떠한가, ② 본고사성적과 1년간의 GPA 각각에 영향을 미치는 설명변수들은 무엇인가, ③ 흔히 알려진 서울의 8학군 효과는 실제로 존재하는가, ④ 고교내신성적을 고교별로 차등점수화 할 필요가 있는가.

연세대학교에서 1994학년에 실시한 특별전형은 GPA로 평가할 때 효과적이었다고 결론을 내릴 수 있다. 수능능력시험과 본고사의 상관계수는 1995학년도 입시에서 인문계열은 0.65, 자연계열은 0.70이었다. 그리고 서울지역이 기타지역보다 높은 상관관계를 보이고 있다. 본고사에 영향을 미치는 설명변수로는 수능능력시험 성적, 고교내신 성적, 성별, 외국어고등학교 출신 여부와 함께 8학군 여부가 포함된다. 반면에 GPA에 영향을 미치는 설명변수군에는 8학군 여부가 포함되지 않는다. 고교내신 성적과 대학 GPA에 대한 상관분석 결과 고교내신은 고교효과가 있음을 알 수 있다. 이는 고교내신의 고교별 차등점수화를 고려할 필요가 있음을 말하여 준다.

I. 서 론

1994학년도 대학입시에서 연세대학교 신촌캠퍼스는 입학정원의 20%를 무시험 특별전형으로 선발하였고, 나머지 80%는 본고사를 실시하는 일반전형으로 선발하

연세대학교 응용통계학과, 서울 서대문구 신촌동 134번지, 120-749.

LG전자, 서울 중구 남대문로5가 537 LG역전 빌딩 판매기획2실 지원팀, 100-095.

였다. 무시험 특별전형은 고교내신 성적이 50%, 수학능력시험 성적이 50%의 비중을 차지하였고, 일반전형은 고교내신 성적 : 수학능력시험 성적 : 본고사 성적 = 40 : 30 : 30의 비중을 차지하였다. 1994학년도 신입생들의 1년간 대학성적(GPA)이 나온 1995년 초에 “과연 1994학년도 입시에서 특별전형제도는 효과적이었는가?”하는 물음에 답을 하기 위하여 본 연구는 시작되었다.

1994학년도 수험생의 입시관련자료와 입학생의 1년간의 GPA자료, 그리고 1995학년도 수험생의 입시관련자료를 다루면서 이러한 자료들로부터 특별전형 효과의 분석 이외에도 다음과 같은 질문들에 대한 답을 유도할 수 있음을 알게 되었다.

- ① 수학능력시험과 대입본고사는 서로 상관이 있는가? 상관이 있다면 그 크기는 얼마인가?
- ② 현행 대학입시제도는 고교내신 성적에 대하여 고교별로 차등 점수화를 하지 못하도록 하고 있다. 그러면, 고교내신 성적은 과연 고교별로 비슷한 효과를 나타내고 있는가?
- ③ 대입 본고사 성적과 대학 1년간의 GPA 각각에 서울의 8학군 효과는 존재하는가?
- ④ 재수, 삼수, 사수한 학생들의 GPA는 고교를 갓 졸업한 학생들의 GPA보다 떨어지는가?
- ⑤ 1994학년도와 1995학년도에 실시하였던 내신 : 수능 : 본고사 = 40 : 30 : 30의 비중은 과연 적절한 가중치 배분인가?
- ⑥ 현재 특별전형과 일반전형으로 운영되는 입시제도는 수험생들이 가지고 있는 예비적 대학 수학능력을 몇 %나 예측할 수 있는 제도인가?

위와 같은 질문은 비단 교육정책가나 교육전문가만이 제기하는 것은 아니다. 대학입시가 온 국민적 관심사였고, 이를 반영이라도 한 듯 입시제도는 지난 10년 동안에도 여러 번의 변경이 있었다. 그러나, 입시정책의 효율성이나 제도변경의 효과를 실증적 분석을 통하여 평가하려는 시도는 흔치 않았다.

본 연구는 연세대학교의 1994·1995학년도 입시자료와 1994학년도 입학생의 1년간 성적(GPA)을 가지고 위에서 열거한 질문에 답을 제시하고자 한다. 이 연구의 자료는 연세대학교 입학생 및 수험생 자료이므로 연구의 결과가 적용되는 범위가 당연히 연세대학교의 경우로 제한되어야 할 것이다.

II. 특별전형의 효과

1994학년도 대학입시에서 연세대학교의 신촌캠퍼스는 20%를 무시험 특차전형으

로, 나머지 80%는 본고사를 실시하는 일반전형으로 신입생을 선발한 바 있다. 1995년 2월 시점에서 특별전형으로 입학한 학생(특별전형군)과 일반전형으로 입학한 학생(일반전형군)의 학업성취도를 비교하고자 한다.

1994학년에 입학한 5,370명의 신입생 중 다음 각 항에 해당하는 학생들은 분석대상에서 제외하였다.

- ① 본고사를 실시하지 않은 매지리 캠퍼스 학생과 예체능계 학생
- ② 특별전형을 실시하지 않은 교육학과 학생
- ③ 예체능 특기생
- ④ 휴학생

분석대상의 학생은 특별전형군과 일반전형군을 합하여 모두 3,531명이었다. 학생들의 1년간 학업성취도는 94-1학기과 94-2학기의 평량평균의 평균(이하 GPA라 칭함)으로서 측정하였다.

학생들의 GPA는 대학 진학 후의 학습능력 이외에도 수강하는 과목의 난이도, 강사의 학점 분포, 학과의 일반적 분위기에 의하여 영향을 받을 수 있다. 또한 학습능력 자체가 고교성적에 영향을 받고, 고교성적은 다시 학과 선정에 중요한 변수로 작용하게 된다. 이러한 여러 요인들은 특별전형군과 일반전형군을 비교하고자 할 때 고려하여야 할 오차요인이라고 할 수 있다.

이상에서 언급한 오차요인을 최소화하는 방법으로서는 단과대학별, 혹은 각 과별로 블럭을 구성하여 두 군을 비교할 수 있다. μ_s 와 μ_g 를 각각 특별전형과 일반전형으로 입학한 학생들의 1994학년도 GPA의 평균이라고 하자. 우리가 검정하고자 하는 가설을 다음과 같이 세울 수 있다.

$$H_0 : \mu_s \leq \mu_g$$

$$H_1 : \mu_s > \mu_g$$

검정방법은 우선 記述統計量을 살펴보고 정규분포의 가정과 등분산 가정이 만족되는지를 조사하였다. 정규분포검정은 SAS¹⁾, PROC UNIVARIATE의 Shapiro-Wilk검정을 사용하였고, 등분산검정은 정규분포 가정에 로버스트(robust)²⁾한 Levene 검정을 적용하였다. 그리고 정규분포검정과 등분산검정은 단과대학별로 특별전형군과 일반전형군 각각에 대하여 실시하였고, 이 검정결과를 그 단과대학에 소속되어 있는 각 과에 적용시켰다. 예를 들어, 문과대학의 특별전형군과 일반전형군이 모두 정규분포 가정을 위반하지만 등분산 가정은 만족한다고 할 때 이 결과를 국문과, 중문과, 영문과, ... 등에도 그대로 적용시켰다. 이는 각 과가 속해 있는 단

1) SAS Institute [6], pp. 617~634 참조.

2) Brown and Forsythe [2] 참조.

과대학이 충분한 표본크기를 갖고 있으며 모집단 분포에 좀더 근접하리라 판단하였기 때문이다.

우선 단과대학별로 두 군을 비교한 결과는 <표 1>에 나타나 있고 이를 그림으로 표시한 것은 <그림 1>에 있다. 모든 단과대학에 대하여 특별전형군 GPA가 일반전형 GPA보다 높으며 그 중 문과, 상경, 이과, 공과, 법과의 경우는 특별전형군 GPA가 일반전형군 GPA보다 유의적으로 높았다. 5개 단과대학 학생수는 2,948명이고 비교대상 학생 3,531명 중 83.5%를 차지하고 있다. 이로 미루어 보다 특별전형군 GPA가 법교적으로 일반전형군 GPA보다 높다는 결론을 유도할 수 있다. <표 1>에

<표 1> 단과대학별 특별전형 입학생과 일반전형 입학생의 GPA 비교

대학	특별전형군			일반전형군			차이	검정 절차 ¹⁾	p값 ²⁾	등분산 검정의 p값 ³⁾	정규성 검정의 p값
	평균	표준 편차	인원	평균	표준 편차	인원					
문과 대학	2.87	0.740	108	2.77	0.700	453	0.10	M-W	0.0412	0.9761	0.0001 ⁴⁾ 0.0001 ⁵⁾
상경 대학	2.64	0.726	123	2.39	0.679	477	0.25	M-W	0.0002	0.5895	0.0050 0.0110
이과 대학	2.55	0.719	123	2.31	0.732	479	0.24	M-W	0.0010	0.2498	0.0415 0.0001
공과 대학	2.51	0.684	210	2.30	0.709	831	0.21	M-W	0.0001	0.5203	0.0285 0.0065
신과 대학	2.86	0.759	9	2.69	0.564	81	0.17	t	0.2007	0.3385	0.1016 0.3882
사회과학 대학	2.78	0.702	47	2.69	0.667	197	0.09	M-W	0.1619	0.6505	0.0148 0.0019
법과 대학	2.84	0.497	30	2.61	0.599	114	0.23	t	0.0298	0.1544	0.5616 0.5907
생활 과학대학	2.95	0.531	29	2.78	0.508	124	0.17	t	0.0514	0.8060	0.3032 0.0951
간호 대학	3.04	0.504	4	2.54	0.538	92	0.50	t	0.0361	0.3201	0.3797 0.2388

1) 정규분포 가정이 만족되면 이표본 t-검정을 사용하고, 그렇지 않으면 맨-휘트니(Mann-Whitney)검정을 사용함.

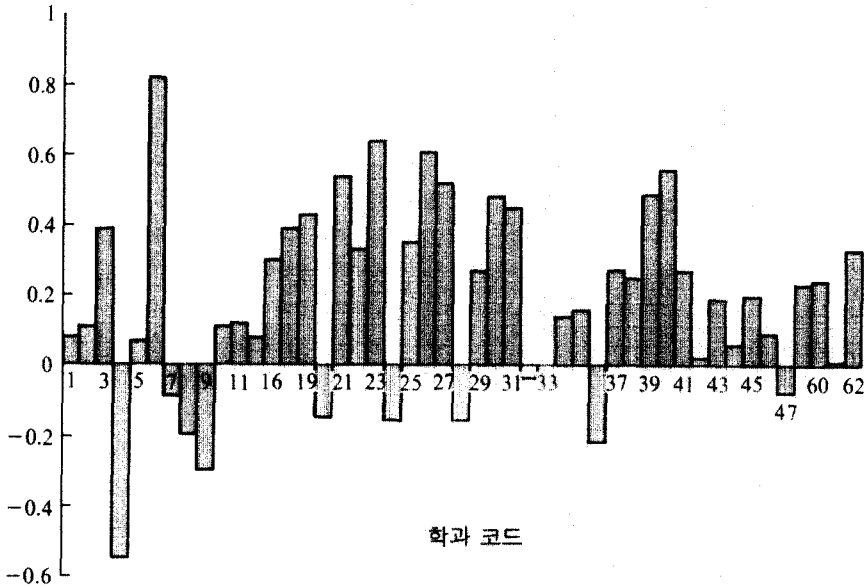
2) 단측검정의 p값임.

3) 르빈(Levene) 검정의 p값임.

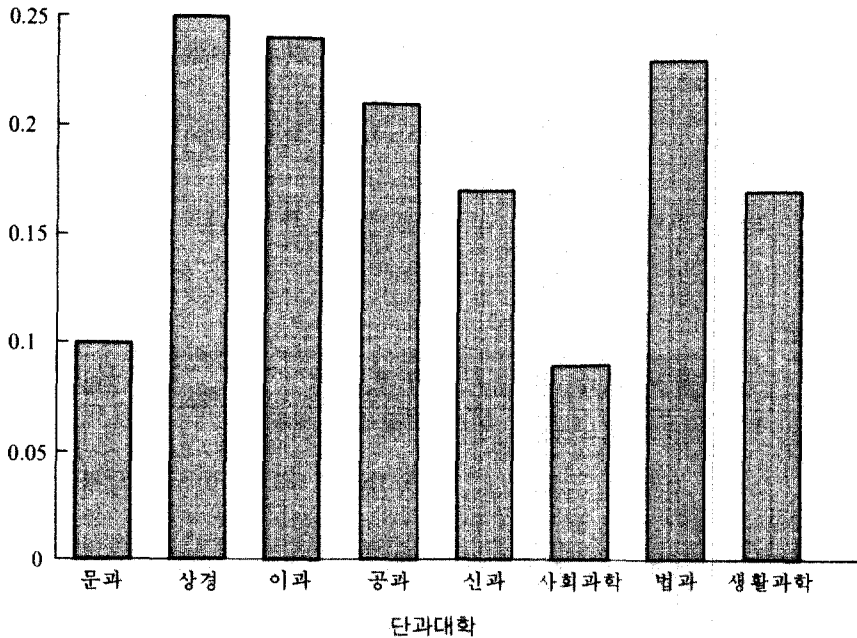
4) 샤피로-윌크(Shapiro-Wilk) 통계량을 이용한 특별전형군의 정규성 검정의 p값임.

5) 샤피로-윌크(Shapiro-Wilk) 통계량을 이용한 일반전형군의 정규성 검정의 p값임.

〈그림 1〉 특차전형 GPA-일반전형 GPA



〈그림 2〉 특차전형 GPA-일반전형 GPA



〈표 2〉 각과별 특별전형 입학생과 일반전형 입학생의 GPA의 비교

문과대학

학과 (학과코드)	특별전형군			일반전형군			차이	검정 절차 ¹⁾	p값 ²⁾
	평균	표준편차	인원	평균	표준편차	인원			
국어국문 (01)	3.10	0.430	12	3.02	0.498	45	0.08	M-W	0.3883
중어중문 (02)	2.75	0.565	8	2.64	0.751	31	0.11	M-W	0.2949
영어영문 (03)	3.11	0.764	19	2.72	0.859	79	0.39	M-W	0.0186
독어독문 (04)	2.49	0.984	9	3.04	0.569	35	-0.55	M-W	0.5
불어불문 (05)	3.05	0.420	7	2.98	0.650	34	0.07	M-W	0.4793
노어노문 (06)	3.60	0.323	8	2.78	0.771	31	0.82	M-W	0.0007
사학 (07)	2.63	0.623	9	2.72	0.633	46	-0.09	M-W	0.5
철학 (08)	2.29	0.955	6	2.49	0.626	32	-0.20	M-W	0.5
문헌정보 (09)	2.35	1.211	7	2.65	0.747	30	-0.30	M-W	0.5
사회 (10)	2.94	0.465	14	2.83	0.618	51	0.11	M-W	0.2616
심리 (11)	2.69	0.555	9	2.57	0.638	39	0.12	M-W	0.4009

상경대학

학과 (학과코드)	특별전형군			일반전형군			차이	검정 절차 ¹⁾	p값 ²⁾
	평균	표준편차	인원	평균	표준편차	인원			
경제 (15)	2.45	0.754	35	2.37	0.710	142	0.08	M-W	0.3200
경영 (16)	2.73	0.694	76	2.43	0.661	289	0.30	M-W	0.0002
응용통계 (17)	2.57	0.790	12	2.18	0.670	46	0.39	M-W	0.0777

이과대학

학과 (학과코드)	특별전형군			일반전형군			차이	검정 절차 ¹⁾	p값 ²⁾
	평균	표준편차	인원	평균	표준편차	인원			
수학 (19)	2.69	0.701	9	2.26	0.733	33	0.43	M-W	0.1572
물리 (20)	2.33	0.627	9	2.48	0.860	39	-0.15	M-W	0.5
화학 (21)	2.92	0.681	11	2.38	0.732	38	0.54	M-W	0.0292
생물 (22)	2.60	0.443	7	2.27	0.586	28	0.33	M-W	0.0559
지질 (23)	2.65	0.503	7	2.01	0.689	28	0.64	M-W	0.0227
천문 (24)	2.11	0.694	6	2.27	0.484	22	-0.16	M-W	0.5
대기 (25)	2.41	0.798	6	2.06	0.966	22	0.35	M-W	0.2598
생화학 (26)	3.03	0.565	8	2.42	0.667	30	0.61	M-W	0.0173
전산과학 (27)	2.79	0.833	16	2.27	0.803	61	0.52	M-W	0.0091
의예 (28)	2.32	0.729	30	2.48	0.723	118	-0.16	M-W	0.5
치의예 (29)	2.39	0.732	14	2.12	0.588	60	0.27	M-W	0.1163

공과대학

학과 (학과코드)	특별전형군			일반전형군			차이	검정 절차 ¹⁾	p값 ²⁾
	평균	표준편차	인원	평균	표준편차	인원			
화학공학 (30)	2.71	0.571	22	2.23	0.679	74	0.48	M-W	0.0018
전기공학 (31)	2.81	0.596	20	2.36	0.707	84	0.45	M-W	0.0049
건축공학 (32)	2.34	0.633	17	2.38	0.608	71	-0.04	M-W	0.5
토목공학 (33)	2.21	0.832	18	2.21	0.794	71	0.00	M-W	0.3642
기계공학 (34)	2.42	0.668	27	2.28	0.685	103	0.14	M-W	0.1315
금속공학 (35)	2.36	0.746	16	2.20	0.720	62	0.16	M-W	0.2384
세라믹공학 (36)	2.28	0.593	13	2.50	0.661	55	-0.22	M-W	0.1374

공과대학

학과 (학과코드)	특별전형군			일반전형군			차이	검정 절차 ¹⁾	p값 ²⁾
	평균	표준편차	인원	평균	표준편차	인원			
전자공학 (37)	2.61	0.664	25	2.34	0.791	105	0.27	M-W	0.1041
식품생물 (38)	2.34	0.736	14	2.09	0.793	51	0.25	M-W	0.1157
전파공학 (39)	2.84	0.476	10	2.35	0.609	39	0.49	M-W	0.0078
기계설계 (40)	2.85	0.846	10	2.29	0.592	39	0.56	M-W	0.0159
도시공학 (41)	2.41	0.476	10	2.14	0.823	38	0.27	M-W	0.2628
산업시스템 (42)	2.45	0.813	8	2.43	0.535	39	0.02	M-W	0.3512

신과대학

학과 (학과코드)	특별전형군			일반전형군			차이	검정 절차 ¹⁾	p값 ²⁾
	평균	표준편차	인원	평균	표준편차	인원			
신학 (43)	2.71	0.76	4	2.52	0.59	47	0.19	t	0.2755
사회사업학 (44)	2.98	0.82	5	2.92	0.44	34	0.06	t	0.3958

사회과학대학

학과 (학과코드)	특별전형군			일반전형군			차이	검정 절차 ¹⁾	p값 ²⁾
	평균	표준편차	인원	평균	표준편차	인원			
정치외교 (45)	2.80	0.655	18	2.60	0.711	74	0.20	M-W	0.1238
행정 (46)	2.82	0.653	18	2.73	0.615	78	0.09	M-W	0.2461
신문방송 (47)	2.69	0.892	11	2.77	0.681	45	-0.08	M-W	0.4345

법과대학

학과 (학과코드)	특별전형군			일반전형군			차이	검정 절차 ¹⁾	p값 ²⁾
	평균	표준편차	인원	평균	표준편차	인원			
법과 (50)	2.84	0.497	30	2.61	0.599	114	0.23	t	0.0298

생활과학대학

학과 (학과코드)	특별전형군			일반전형군			차이	검정 절차 ¹⁾	p값 ²⁾
	평균	표준편차	인원	평균	표준편차	인원			
의류환경 (60)	3.13	0.541	8	2.89	0.504	31	0.24	t	0.1259
식품영양 (61)	2.69	0.705	7	2.68	0.527	30	0.01	t	0.4831
주거환경 (62)	3.07	0.317	8	2.74	0.544	30	0.33	t	0.0524
아 동 학 (63)	2.85	0.496	6	2.79	0.458	33	0.06	t	0.3839

간호대학

학과 (학과코드)	특별전형군			일반전형군			차이	검정 절차 ¹⁾	p값 ²⁾
	평균	표준편차	인원	평균	표준편차	인원			
간호학 (70)	3.04	0.504	4	2.54	0.538	92	0.50	t	0.0361

1) 각 과가 소속되어 있는 단과대학의 GPA 분포가 정규성을 만족하면 이표본 t-검정을, 그렇지 않으면 맨-휘트니검정을 실시함. <표 1>에서 보듯이 모든 단과대학에 걸쳐서 두 군의 등분산 가정은 만족되었으므로, 모든 과에 대하여 등분산이 성립한다고 가정함.

2) 맨-휘트니검정(M-W)이나 이표본 t-검정(t)에 따른 단측검정의 p값임.

서 간호대학의 경우는 특별전형의 학생수가 4명으로 극히 적은 숫자이다. 따라서 간호대학의 경우 특별전형군과 일반전형군의 비교에 대하여서는 결론을 유보하고자 한다.

각 학과별로 두 군을 비교한 결과는 <표 2>에 나타나 있고, 이를 그림으로 표시한 것은 <그림 2>에 있다. 비교 대상인 49개 학과 중 39개 학과에서 특별전형군 GPA가 더 높으며, 그 중 12개 학과에 대하여서는 그 차이가 통계적으로 유의하였다.

이상의 분석 결과는 특별전형군 GPA가 일반전형군 GPA보다 높다는 증거를 제시하여 준다.

Ⅲ. 수학능력시험과 본고사 성적 간의 상관관계

본고사가 부활된 지난 몇 년 동안 대학입시에서 본고사가 과연 필요한가에 대한 논란이 있었고 아직도 이에 대한 논란은 계속되고 있다. 수학능력시험이 본고사를 대체할 수 있는가 하는 것을 규명하기 위하여서는 우선 두 시험간의 상관관계를 파악하여야 한다. 본 분석에서는 1994학년도와 1995학년도 입시자료에 기초하여 두 시험성적간의 상관계수를 계산하고자 한다. 수학능력시험 성적과 본고사 성적 간의 상관분석을 하는데 흔히 범할 수 있는 오류는 합격생들만을 분석대상으로 한다는 점이다. 본고사를 치르는 대학의 합격선은 본고사 점수와 수학능력시험 점수가 서로 보완적으로 작용하여 결정된다. 즉, 수학능력시험 점수가 다소 낮더라도 본고사 점수가 높으면 합격이 가능하다. 따라서 합격생들만을 대상으로 하여 계산한 두 시험 성적간의 상관계수는 필연적으로 낮을 수밖에 없고, 이는 두 시험 성적간의 전반적인 추세를 반영하고 있지 않다. 두 시험 성적간의 상관분석은 지원자들을 분석대상으로 하여야 한다는 점을 강조하고자 한다.

1. 1994학년도 입시자료 분석

1994학년도 대학입시에서 연세대학교는 입학정원의 20%를 특별전형으로, 나머지 80%는 본고사를 실시하는 일반전형으로 신입생을 선발한 바 있다. 본절에서는 우선 1994학년도 자료에 기초하여 수학능력 성적과 본고사 성적 간의 상관분석을 실시하고, 현행 입시와 관련되어 제기할 수 있는 가설들을 검정하고자 한다.

본 논의에서는 상관관계를 지역별, 계열별로 비교 분석하고자 한다. 이는 서울 지역 학생들은 기타지역, 특히 소도시나 군지역 학생들보다 본고사 관련 정보를 더 많이 접할 수 있으므로 그에 따른 효과가 높은 상관관계로 나타나리라 기대하기 때문이다. 또한 인문, 자연 두 계열간에는 시험과목의 차이로 인하여 분포의 이질성이 존재하리라 사료된다.

ρ_S 와 ρ_{NS} 를 각각 서울과 그 외의 도시의 본고사 점수와 수학능력시험(이하 수능) 점수간의 상관관계라고 하자. 검정하고자 하는 가설은 다음과 같다.

$$H_0 : \rho_S \leq \rho_{NS}$$

$$H_1 : \rho_S > \rho_{NS}$$

즉, '서울지역 학생들이 비서울지역 학생들에 비하여 본고사 성적과 수학능력시험 성적과의 상관관계가 더 높다'를 제시하려는 가설로 세우고, 두 상관관계의 차

에 대한 검정 절차로서 피셔의 Z변환(Fisher's Z)³⁾을 이용한 방법을 사용한다. 위와 같은 가설 설정은 서울학생들이 타도시의 학생들에 비해 본고사에 대비한 모의시험과 정보 등이 많음으로 인하여 서울지역 이외 학생들에 비하여 같은 수능 점수의 수준에서 본고사 점수의 분포가 상대적으로 밀집되어 있을 것이라는 가정에 기초한다. 또한 서울에서도 8학군에 해당하는 학생들은 과외와 학원 수강 등으로 수능 시험 이후 본고사를 대비함으로써 본고사의 성적이 타지역의 같은 수능 점수 학생들에 비해 상대적으로 높으리라 예상할 수 있다. 따라서 8학군 학생들은 낮은 수능 성적을 보완할 수 있을 만한 높은 본고사 성적을 얻음으로써 상관계수가 다소 낮으리라 사료된다. 따라서 서울지역의 효과, 더 구체적으로 8학군의 효과를 검정하기 위하여 지역을 다음과 같이 분류하였다. 즉, 전국을 서울의 8학군, 서울에서 8학군을 제외한 지역, 기타도시로 분류하였다. ρ_0 과 ρ_8 를 각각 서울의 8학군 지역과 그 이외 지역의 수능과 본고사의 상관관계로 표시할 때 다음과 같은 가설을 세웠다.

$$H_0 : \rho_0 \leq \rho_8$$

$$H_1 : \rho_0 > \rho_8$$

마지막으로 계열을 불력으로 구성하여 수능과 본고사 점수의 상관관계를 비교 분석하고자 한다. ρ_N , ρ_H 각각을 자연계열과 인문계열의 본고사와 수능 점수 간의 상관관계라 나타내고 다음과 같은 가설을 세우고자 한다.

$$H_0 : \rho_N = \rho_H$$

$$H_1 : \rho_N \neq \rho_H$$

수능 점수와 본고사 점수의 상관관계를 계산하는 과정에서 신촌 캠퍼스 중에서 실기전형울 실시한 음악대학과 사회체육과, 체육교육과를 제외하였고, 본고사를 치르지 않은 특별전형의 학생들도 이 계산에서 제외하였다.

위에서 언급한 내용의 상관관계는 <표 3>에 나타나 있다. 8학군지역은 낮은 상관관계를, 서울지역이 비서울지역보다는 높은 상관관계를 보였고, 인문계열보다는 자연계열이 더 높은 상관관계를 보였다.

분석결과 현행 입시제도하에서 주목할 만한 사항은 다음과 같다.

- ① 1994학년도 입시자료에서는 인문계열과 자연계열을 분리하여 상관계수를 계산하는 것이 타당하다. <표 3>에서 보듯이 인문계열과 자연계열 각각의 상관계수가 0.72와 0.77임에도 전체의 상관계수가 흔히 기대할 수 있는 0.72와 0.77 사이의 값이 아닌 0.59를 나타내는 것은 두 계열간 분포의 이질성을 나타내 준다고 할 수 있다. 1994학년도 자연계열의 본고사 성적과 수능 성적의 점산도인

3) Anderson [1], pp. 78~79 참조.

〈표 3〉 1994학년도 수학능력 성적과 본고사 성적의 상관분석표

	서울	기타 도시	p값 ¹⁾	전체
수능과 본고사 성적의 상관관계	0.60 ²⁾ (2958)	0.55 (1938)	0.005	0.59 (4896)

	8 학 군	서울-8학군 ³⁾	기타지역	비 고	전 체
수능과 본고사 성적의 상관관계	0.51 (921)	0.63 (2037)	0.55 (1938)	8학군과 기타지역은 비유의적이고, 서울-8학군과 기타지역은 유의적임 ⁴⁾	0.59 (4896)

	인문계열	자연계열	p값	전체
수능과 본고사 성적의 상관관계	0.72 (2521)	0.77 (2375)	0.0000	0.59 (4896)

- 1) 두 상관계수의 차이에 대한 유의성 검정임.
- 2) 상관계수 밑의 () 안의 숫자는 대상 인원수를 나타냄.
- 3) 서울-8학군은 서울지역에서 8학군지역을 제외한 지역을 나타냄.
- 4) 유의수준 0.05에서 본페로니(Bonferoni)검정을 실시함.

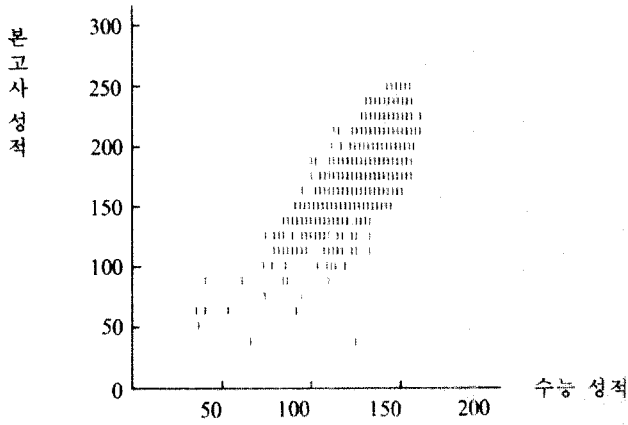
〈그림 4〉와 인문계열의 점산도인 〈그림 3〉을 비교하여 보면 〈그림 4〉는 〈그림 3〉을 오른쪽으로 수평이동한 형태임을 알 수 있고, 〈그림 5〉에서 보듯이 두 점산도를 합하면 더욱 '블록'한 형태가 된다. 따라서 합한 점산도의 상관계수가 어느 한 점산도가 갖는 상관계수보다 더 작은 값이 됨을 알 수 있다.

- ② 〈표 3〉에서 보듯이 1994학년도 서울지역의 상관계수가 기타 지역보다 높은 것은 서울지역 학생들이 각종 모의고사등을 통하여 본고사시험 준비 기회를 더 많이 가질 수 있고, 따라서 같은 수능 성적의 학생들이라도 서울지역 학생들의 본고사 성적의 분산이 기타지역 학생들의 분산보다 작으리라 판단된다. 이러한 현상이 서울지역의 높은 상관계수로 반영된다고 사료된다. 서울지역 효과에 대한 좀더 세밀한 분석으로서 서울의 8학군 지역 학생들의 본고사 성적이 같은 수능 성적 분포군의 기타지역에 비해 높을 것이라 추측할 수는 있겠으나 그 효과가 상관분석으로는 측정되지 않는다.
- ③ 1994학년도 입시에서 인문계열보다는 자연계열의 수능 성적과 본고사 성적의 상관계수가 더 높다.

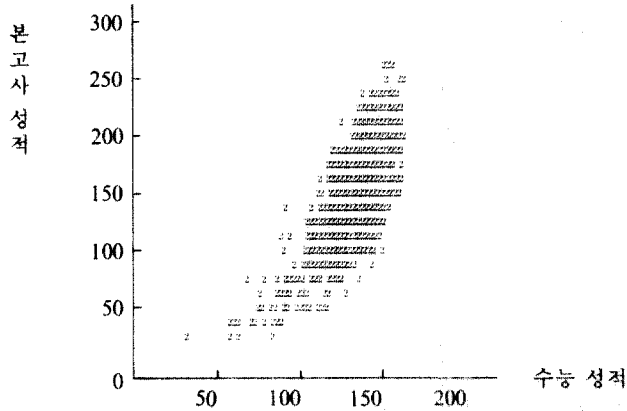
2. 1995학년도 자료 분석

1995학년도 대학입시에서 연세대학교는 입학정원의 40%를 특별전형으로, 나머지 60%는 본고사를 실시하는 일반전형으로 신입생을 선발한 바 있다. 1995학년도

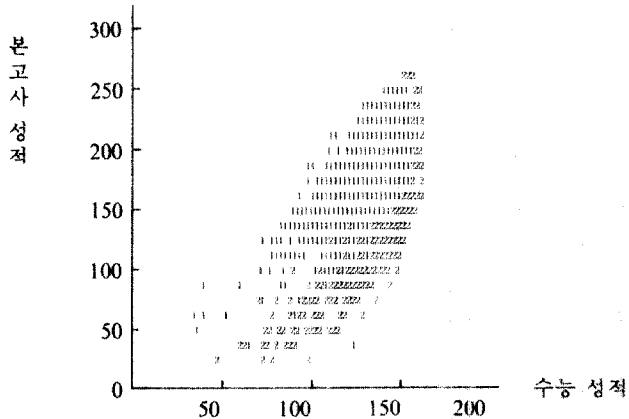
<그림 3> 1994학년도 인문계열의 수학능력 성적과 본고사 성적의 점산도



<그림 4> 1994학년도 자연계열의 수학능력 성적과 본고사 성적의 점산도



<그림 5> 1994학년도 수학능력 성적과 본고사 성적의 점산도(전체수험생 대상)



입시자료를 이용하여 지역과 계열에 따른 수능과 본고사의 상관계수를 구하고, 1995학년도 입시에서 나타난 외국어고등학교 효과가 수능과 본고사의 상관관계에 어떻게 작용하는지를 알아본다.

신촌 캠퍼스 중에서 실기전형을 실시한 음악대학과 사회체육과, 체육교육과를 제외하였으며, 본고사를 치르지 않은 특별전형의 학생들과 원주 캠퍼스도 본 분석에서 제외하였다. 1995학년도부터 실시된 외국어고등학교 출신자의 어문계열 진학 시 받는 내신 특례로 인하여 1995학년도 입시자료 분석에서 외국어고등학교의 특수성을 고려해 볼 수 있다. 외국어고등학교 출신자들이 많은 어문학 계열을 비어문학 계열과 비교하여 수능과 본고사의 상관계수를 구하고자 한다.

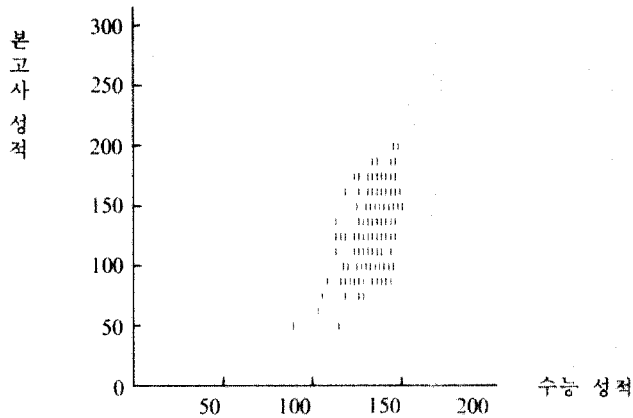
외고 출신자들은 성적 분포가 비교적 동질적인 그룹으로 일정 점수대에 밀집되어 있을 것이라 추측된다. 따라서 비어문학 계열 학생들의 본고사와 수능의 상관관계보다는 낮게 되리라 사료된다. 이에 대한 분석으로서 어문학과, 비어문학과로 분류집단을 나누어 분석해 보았다. 또한 1994학년도에는 자연계열이 인문계열에 비해 상대적으로 수능과 본고사 성적에 있어 높은 상관관계를 나타내었는데, 이와 같은 현상이 1995학년도에도 지속되는지를 조사하기 위하여 자연계열과 인문계열별로 수능과 본고사점수 간에 상관관계를 계산하였다. 상관분석의 결과는 <표 4>에

<표 4> 1995학년도 수능능력 성적과 본고사 성적의 상관분석표

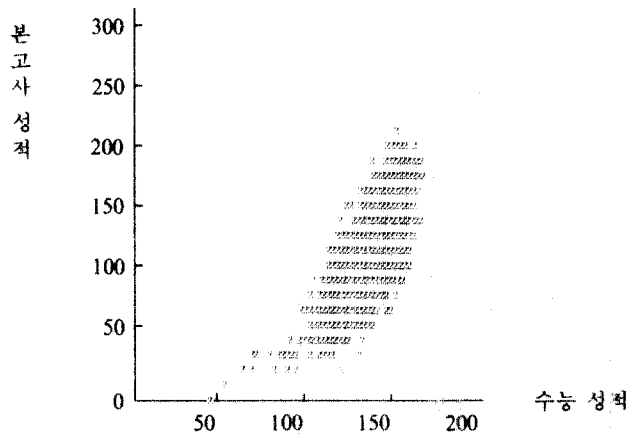
	서울	기타 도시	p 값 ¹⁾	전체	
수능과 본고사 성적의 상관관계	0.69 ²⁾ (3138)	0.62 (2160)	0.0000	0.67 (5298)	
	8학군	서울-8학군 ³⁾	기타지역	비교	전체
수능과 본고사 성적의 상관관계	0.67 (979)	0.69 (2159)	0.62 (2160)	8학군지역과 서울-8학군은 모두 기타지역은 유의적임 ⁴⁾	0.67 (5298)
	어문학계열	비어문학계열	p 값	전체	
수능과 본고사 성적의 상관관계	0.51 (310)	0.68 (4988)	0.0000	0.67 (5298)	
	인문계열	자연계열	p 값	전체	
수능과 본고사 성적의 상관관계	0.65 (2669)	0.70 (2629)	0.0004	0.67 (5298)	

- 1) 두 상관계수의 차이에 대한 유의성 검정임.
- 2) 상관계수 밑의 () 안의 숫자는 대상 인원을 나타냄.
- 3) 서울-8학군은 서울지역에서 8학군 지역을 제외한 지역을 나타냄.
- 4) 유의수준 0.05에서 본페로니(Bonferoni)검정을 실시함.

〈그림 6〉 1995학년도 어문학계열의 수능능력 성적과 본고사 성적의 점산도



〈그림 7〉 1995학년도 비어문학계열의 수능능력 성적과 본고사 성적의 점산도



나타나 있다.

〈표 4〉에서 보듯이 지역별 효과에 대한 1995학년도 입시자료 분석의 결과는 1994학년도와 비슷하였다. 즉, 계열별 구별이 없이 전 수험생을 대상으로 한 본고사와 수능 성적 간의 상관계수는 0.67로 계산되었고 서울지역이 0.69, 비서울지역이 0.62로 서울지역의 상관계수가 좀더 높은 것으로 계산되었다. 서울의 8학군지역은 0.67, 8학군을 제외한 서울지역은 0.69로 상관관계에 있어서는 별다른 차이를 보이지 않았다. 1994학년도 입시자료에서 8학군지역의 상관계수가 낮게 나온 것과는 달리 1995학년도 입시자료에서는 〈표 4〉에서 보듯이 8학군지역의 상관계수가 8학군을 제외한 서울지역(서울-8학군)과 큰 차이를 보이지 않고 있다.

1995학년도 입시자료 분석에서 주목할 만한 사항은 첫째, 1995학년도 입시에서

외국어고등학교의 내신 특례로 인한 외고효과를 분석해 본 결과 외고출신이 많은 어문학계열은 수능과 본고사에 있어 낮은 상관관계를 보여 주고 있다. 이는 <그림 6>에서 보듯이 외국어고등학교 출신자가 많은 어문학계열 학생들의 성적 분포가 <그림 7>의 비어문학 계열 학생들의 성적분포에 비하여 일정 점수대에 밀집되어 있어 본고사와 수능 성적의 낮은 상관관계를 보인다. 둘째, 인문계열보다는 자연계열이 더 높은 상관관계를 보였다. 셋째, 1994학년도와 마찬가지로 서울지역의 상관계수가 기타 지역보다 높게 나타났다.

IV. 1995학년도 입시 본고사 성적의 분석

1995학년도 본고사 성적에 영향을 미치는 요인을 알아보고 본고사와 관련한 가설, 예를 들어 8학군 효과등을 검정하기로 한다. 일반적으로 다중회귀분석은 반응변수 Y 와 p 개의 설명변수 X_1, \dots, X_p 의 관계를 다음 식 (1)과 같이 선형함수로 구성한다.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \quad (1)$$

단, ε 은 p 개의 설명변수로서 나타낼 수 없는 부분들을 포괄적으로 표시하는 오차항이고, 평균이 0이고 일정한 분산을 갖는 정규분포를 이룬다고 가정한다.

본 분석에서는 1995학년도 입시에서 일반전형을 지원한 수험생 중 다음 어느 하나에라도 해당되는 수험생은 제외하였다.

- ① 예체능 특기생
- ② 본고사 결시자 혹은 면접결시자
- ③ 이중지원등의 미자격자
- ④ 본고사 시험을 치루지 않은 특별전형 학생
- ⑤ 수능 성적은 높으나 본고사 성적이 10점 미만인자

②의 수험생은 본고사전형을 포기한 것으로 간주하였고 수험생 2명은 이상치로 판단되어 분석에서 제외하였다. ①~⑤에 해당되는 수험생을 제외하면 분석대상 수험생은 모두 5,296명이 된다. 반응변수인 본고사 성적(BON)의 가능한 설명변수들로서 <표 5>의 변수들을 고려하였다.

<표 5>의 변수들 중 *Group8*과 *Seoul_8*은 가변수로서 전국의 고등학교를 <표 6>과 같이 3개 지역으로 구분하기 위하여 만들어졌다.

반응변수로서 변환하지 않은 본고사 성적, 즉 BON을 사용하여 모형을 추정하고

잔차도를 그려 보았더니 본고사 성적이 좋은 학생들의 분포가 오히려 성적이 낮은 학생들에 비하여 분산이 더 큰 것을 볼 수 있었다. 따라서 성적이 낮은 쪽의 값을 상대적으로 더 크게 하고 높은 쪽은 작게 하는 로그변환을 취해 보았다. 나팔과 같은 형태의 잔차도는 반응변수의 로그변환이나 제곱근 변환을 함으로써 정상적인 잔차도를 얻을 수 있는 대표적인 경우이다.⁴⁾ 이 밖에도 반응변수인 본고사 성적 (*BON*)에다 로그변환과 제곱근변환을 모두 시도하여 본 결과 로그변환이 더 좋은 적합도를 나타내었으며, 높은 R^2 를 나타내었다. 따라서 반응변수로는 $\text{LOG}(BON)$ 를 사용하였다. 또한 본고사 성적의 다중회귀분석을 위한 설명변수로서 수능 성적

〈표 5〉 본고사 성적의 가능한 설명변수들

설 명 변 수	영 문 변 수 명
수능 성적	<i>SAT</i>
내신 백분율	<i>NPER</i>
성별	<i>SEX</i> (1: 남자, 0: 여자)
어문계열 여부	<i>LANG</i> (1: 어문, 0: 비어문)
인문·자연계열 구분	<i>ARTS</i> (1: 인문, 0: 자연)
외고 출신 여부	<i>FOR_LAN</i> (1: 외고, 0: 비외고)
서울의 8학군 여부	<i>Group 8</i> (1: 8학군, 0: 8학군 아님)
8학군의 서울지역 여부	<i>Seoul_8</i> (1: 8학군이 아닌 서울지역, 0: 기타지역)

〈표 6〉 전국고등학교의 지역구분

지 역
서울의 8학군
서울지역에서 8학군을 제외한 지역
서울 이외의 지역

〈표 7〉 최종모형의 모수 추정결과

반응변수	설명변수	회귀계수	t 통계량값	p 값
$\text{LOG}(BON)$	절편항	3.84	198.502	0.0000
	<i>SAT**2</i>	0.000043	58.353	0.0000
	<i>NPER</i>	-0.003726	-21.230	0.0001
	<i>SEX</i>	-0.031790	-5.838	0.0001
	<i>ARTS</i>	0.035130	7.319	0.0001
	<i>FOR_LAN</i>	0.094526	6.244	0.0001
	<i>Group8</i>	0.061332	9.986	0.0001

$R^2 = 0.537(\bar{R}^2 = 0.536)$, Mallows' $C_p = 8.6$.

4) Neter, Wasserman and Kutner [3], pp. 116~130.

을 제공하였다. 이는 점산도를 통하여 관찰할 때 본고사 성적과 수능 성적이 선형적 관계가 아닌 곡선의 형태를 보이고 있기 때문이다.

회귀분석에서 최적 설명변수군을 찾는 방법 중 하나인 후진제거법(Backward Elimination)을 사용하여 얻은 최종모형의 추정결과는 <표 7>과 같다.

<표 7>의 추정결과로부터 다음의 사실을 도출할 수 있다.

- ① 흔히 생각할 수 있는 설명변수인 수능 성적, 내신 백분비, 인문·자연 구분 이외에 성별 효과, 외고 효과, 그리고 8학군 효과가 유의적임이 확인되었다.
- ② 다른 모든 조건이 같다면 외고 효과는 내신 백분비로 환산하여 25.40%를 향상시키는 효과와 같다. 예를 들어, 외고에서 내신 7등급은 비외고의 내신 3등급~4등급에 해당한다.
- ③ 다른 조건이 다 같은 경우 수능 성적 150점과 160점 간의 차이가 본고사에 미치는 영향은 내신 백분비 35.77%의 차이가 본고사에 미치는 영향과 같다고 할 수 있다.

V. 1994학년도 입학생 GPA의 다중회귀분석

대학학업성취도를 나타내는 GPA를 반응변수로 하는 다중회귀분석 모형을 식 (1)과 유사하게 구성하여 적절한 설명변수를 구성하고 반응변수의 변환등을 시도하여 최적 모형을 찾고자 한다. 본 분석에서는 1994학년도에 입학한 5,370명 학생 중 예체능계 학생, 원주 캠퍼스 학생, 예체능 특기생과 휴학생을 제외한 3,588명을 대상으로 한다. 반응변수인 GPA에 영향을 미치리라 판단되는 설명변수들로서 <표 8>의 변수를 고려하였다. <표 8>에서 일반전형은 치룬 학생에 한하여 본고사 성적을 설명변수로 사용할 수 있으므로, 이를 나타내기 위하여 BON2를 식 (2)와 같이 정의하였다.

$$BON2 = (1 - SG) \times BON \quad (2)$$

(단, BON은 본고사 성적이고 SG = 1은 특별전형,

SG = 0은 일반전형을 의미한다)

<표 8>에서 i 학과는 본교의 학과코드를 따른다. 예를 들어, 국문학과는 Dept 1, 수학과 Dept 19 등으로 나타내고, 본고사를 치룬 50개 학과를 나타내기 위하여 49개의 가변수를 사용하였다. 8학군 여부를 나타내는 Group 8과 서울지역 여부를 나타내는 SEOUL은 모형에 동시에 포함시키지 않고 하나씩 포함시켜 적합도가 좋은 지역변수를 찾아 보았다. 분석대상이 되는 3,588명 중 특수고 출신은 모두 24명으

〈표 8〉 GPA의 가능한 설명변수들

설명변수	영문변수명
특차전형 여부	SG(1: 특차, 0: 일반)
본고사 성적	BON2 = (1 - SG) × BON
수능 성적	SAT
내신 백분율	NPER
성별	SEX(1: 남자, 0: 여자)
재수 여부	RETRY(1: 재수 이상, 0: 현역)
i 학과 구분	Dept <i>i</i> (<i>i</i> = 1, ..., 49)
특수고 출신 여부	SPECIAL(1: 특수고, 0: 기타고교)
외고 출신 여부	FOR_LAN(1: 외고 출신, 0: 비외고 출신)
서울 8학군 여부	Group8(1: 8학군지역, 0: 8학군 아님)
서울지역 여부	SEOUL(1: 서울지역, 0: 기타지역)

로 많지 않았으나, 1994학년도 입시에서 내신특례가 적용된 과기고와 청주외고 등으로 구성되는 특수고 학생들의 우수성이 대학입학 후에도 지속되는지를 알아보기 위하여 잠재적 설명변수로 고려하였다.

반응변수 GPA를 변환하여 GPA²로 다시 회귀분석을 해 본 결과 회귀분석 모형의 잔차도 분석 결과 GPA²가 통계학적으로 더 바람직한 잔차도를 가짐을 알 수 있었다. 따라서 GPA²를 반응변수로 하여 여러 개의 다중회귀분석 모형을 구성하였다. GPA²를 반응변수로 하여 여러 개의 다중회귀분석 모형을 구성하고 추정된 결과 〈표 9〉 모형을 최적 모형으로 판단하였다.

〈표 9〉의 최종 모형에서 주목할 만한 점들은 다음과 같다.

- ① 8학군 효과가 1995학년도 본고사 성적에는 유의적인 설명변수로 나타나지만, GPA²에 대한 설명변수로는 적합치 않다는 사실은 8학군 효과가 입학한 1년 후에는 소멸한다는 것을 간접적으로 증거한다고 말할 수 있다.
- ② 특수고 출신 여부는 GPA²에 대하여 유의적인 설명변수로 나타났다. 1994학년도 입학생 중 특수고 출신은 24명밖에 되지 않으나 그 중 21명이 각 소속과의 평균 GPA보다 높은 성적을 나타내고 있다.
- ③ 재수 혹은 삼수, 사수는 대학학업성취도에 역효과를 초래한다. 즉, 재수 이상을 하여 대학에 입학한 학생들이 갓 졸업한 학생들보다 통계학적으로 유의하게 GPA면에서 떨어졌다.
- ④ 학과효과가 존재한다. 즉, 학과에 따라 GPA²의 평균의 차이가 있고, 이러한 학과효과는 상당히 많은 학과에서 나타나고 있고, 학과효과는 각 과의 계수의 크기를 비교함으로써 알 수 있다. 예를 들어, 국문학과와 경우 Dept 1의 계수가

〈표 9〉 반응변수 GPA의 최적 설명변수들

설명변수	회귀계수	t 통계량값	p 값
절편	-0.150	-0.825	0.4095
특별전형여부	7.419	11.252	0.0001
수능 성적	0.041	4.782	0.0001
내신 백분비	-0.065	-12.257	0.0001
성별(1:남자, 0:여자)	-2.289	-16.188	0.0001
본고사 성적	0.038	11.616	0.0001
재수 여부	-0.488	-4.143	0.0001
특수고 출신	1.855	2.969	0.0030
국문학과	-2.706	-4.305	0.0001
중문학과	-4.552	-6.624	0.0001
영문학과	-4.339	-7.091	0.0001
독문학과	-2.362	-3.606	0.0003
불문학과	-3.122	-4.615	0.0001
노문학과	-1.960	-2.914	0.0036
사학과	-3.387	-5.403	0.0001
철학과	-3.943	-5.729	0.0001
문헌정보학과	-3.919	-5.700	0.0001
사회학과	-3.118	-4.980	0.0001
심리학과	-5.052	-7.546	0.0001
경제학과	-4.716	-8.225	0.0001
경영학과	-4.404	-7.924	0.0001
응용통계학과	-5.788	-8.982	0.0001
수학과	-4.362	-6.696	0.0001
물리학과	-3.365	-5.233	0.0001
화학과	-3.671	-5.958	0.0001
생물학과	-3.916	-5.949	0.0001
지질학과	-3.081	-4.836	0.0001
천문학 전공	-3.667	-5.128	0.0001
대기학 전공	-3.029	-4.435	0.0001
생화학과	-4.079	-6.038	0.0001
전산과학과	-4.391	-7.288	0.0001
의예과	-5.531	-8.541	0.0001
치의예과	-6.379	-9.686	0.0001
화학공학과	-4.128	-7.322	0.0001
전기공학과	-3.432	-5.989	0.0001
건축공학과	-4.842	-7.948	0.0001

설명변수	회귀계수	t 통계량값	p 값
토목공학과	-3.939	-6.984	0.0001
기계공학과	-4.411	-7.745	0.0001
금속공학과	-3.919	-6.797	0.0001
세라믹공학과	-3.244	-5.489	0.0001
전자공학과	-4.322	-7.370	0.0001
식품생물공학과	-4.302	-7.596	0.0001
전파공학과	-4.070	-6.107	0.0001
기계설계학과	-3.865	-5.862	0.0001
도시공학과	-5.069	-7.749	0.0001
산업시스템공학과	-3.900	-5.916	0.0001
신학과	-1.251	-2.159	0.0309
사회사업학과	-1.883	-2.872	0.0041
정치외교학과	-3.858	-6.343	0.0001
행정학과	-3.448	-5.695	0.0001
신문방송학과	-3.723	-5.710	0.0001
법학과	-4.038	-7.049	0.0001
의생활학과	-1.586	-2.549	0.0109
식품영양학과	-1.860	-3.074	0.0021
주생활학과	-2.480	-3.867	0.0001
아동학과	-3.760	-5.737	0.0001
교육학과	-4.584	-7.509	0.0001

$R^2 = 0.250(\bar{R}^2 = 0.238)$

-2.706이고 경제학과의 경우 Dept 15의 계수가 -4.716이므로 국문학과 학생들의 GPA가 경제학과 학생들의 GPA보다 평균 2.01이 높으며, GPA로 환산하면 평균 1.42가 높다.

- ⑤ 수능성적, 내신 백분비, 본고사 성적, 성별이 유의적인 설명변수로 작용하고 있으며 수능 성적과 본고사 성적 중 어느 것이 반응변수에 더 큰 영향을 미치는지는 좀더 연구가 필요하다고 본다.
- ⑥ 특별전형 효과는 본고사 195.2에 해당하는 효과를 나타내고 있다. 즉, 특별전형 학생의 GPA는 본고사 성적을 195.2 받은 학생의 GPA와 같다고 할 수 있다. 또한 예를 들어 특별전형으로 영문학과에 입학한, 재수를 하지 않은, 수능 173점, 내신 3.5%의 여학생은 평균 GPA가 3.64이고, 일반전형으로 영문학과에 입학한, 재수를 하지 않은, 수능 165점, 본고사 200점, 내신 3.5%의 여학생은 평균 GPA가 3.62가 된다.

⑦ 내신, 수능 성적, 그리고 본고사 성적이 대학학업성취도에 미치는 정도를 다음과 같이 비교해 볼 수 있다.

내신 3등급과 2등급의 차이(평균 내신 백분비 4.5%의 차이와 같음)는 수능 성적 7.13점 차이와 같으며, 본고사 점수의 7.70점 차이와 같다. 일반전형 입시의 1000점 만점 중 각각의 차이가 차지하는 비중은 5점, 10.70점, 그리고 7.70점이 된다. 비슷한 논리로 내신 4등급과 3등급의 차이(평균 내신 백분비 5.5%의 차이)는 수능 성적 8.72점 차이와 같으며, 본고사 점수의 9.41점 차이와 같다. 일반전형 입시의 1000점 만점 중 각 성분이 차지하는 비중은 각각 5점, 13.08점, 그리고 7.70점이 된다. 내신 4등급과 3등급의 차이, 수능 성적 8.72점의 차이, 본고사 점수 9.41점의 차이는 모두 같은 크기의 GPA의 차이를 결과하므로, 이상적으로는 일반전형 입시의 1000점 만점 중 이 세 개의 값들의 차이도 비슷한 크기를 지녀야 할 것이다. 그러나 현행 내신 : 수능 : 본고사 = 40 : 30 : 30의 비중에서는 두 배 이상의 크기의 차이가 있다. 따라서 이에 대하여 더 세부적인 분석을 하기로 한다.

우선 분석대상이 된 1994 입학생 중 3,588명의 내신등급의 분포는 <표 10>과 같다. 전체의 85.8%를 차지하는 내신 1등급에서 내신 4등급까지의 학생을 대상으로 내신 등급간의 차이가 GPA에 미치는 효과와 같은 효과를 미치는 수능성적의 차이와 본고사 성적의 차이를 입시 총점 1000점 만점에서 차지하는 크기로써 비교하기로 한다. 내신 : 수능 : 본고사의 비중을 현재의 40 : 30 : 30과 가상적으로 45 : 25 : 30,

<표 10> 분석대상 학생의 내신등급 분포

내신등급	뒸 수	%	누적뒸수	누적%
1	1018	28.4	1018	28.4
2	989	27.6	2007	55.9
3	681	19.0	2688	74.9
4	390	10.9	3078	85.8
5	224	6.2	3302	92.0
6	98	2.7	3400	94.8
7	89	2.5	3489	97.2
8	60	1.7	3549	98.9
9	20	0.6	3569	99.5
10	10	0.3	3579	99.7
11	4	0.1	3583	99.9
12	2	0.1	3585	99.9
13	2	0.1	3587	100.0
14	1	0.0	3588	100.0

〈표 11〉 내신 1등급과 2등급 차이가 GPA에 미치는 효과와 같은 효과를 미치는 수능 점수의 차이와 본고사 점수의 차이
(1000점 만점 중 차지하는 점수)

내신 : 수능 : 본고사의 비중	내신 1등급-내신 2등급 (평균내신백분비 3.5% 차이)	내신 1등급-내신 2등급과 같은 효과를 내는 수능 점수 ()안은 200점 만점의 점수	내신 1등급-내신 2등급과 같은 효과를 내는 수능 점수 ()안은 300점 만점의 점수
40 : 30 : 30 1000점 중 점수 회귀계수	5 0.065	8.32(5.55) 0.041	5.99(5.99) 0.038
45 : 25 : 30 1000점 중 점수 회귀계수	5.63 0.058	5.16(4.13) 0.0492	5.34(5.34) 0.038
50 : 20 : 30 1000점 중 점수 회귀계수	6.25 0.052	2.96(2.96) 0.0615	4.79(4.79) 0.038
50 : 25 : 25 1000점 중 점수 회귀계수	6.25 0.052	4.625(3.70) 0.0492	3.325(3.99) 0.0456

〈표 12〉 내신 2등급과 3등급 차이가 GPA에 미치는 효과와 같은 효과를 미치는 수능 점수의 차이와 본고사 점수의 차이
(1000점 만점 중 차지하는 점수)

내신 : 수능 : 본고사의 비중		내신 2등급-내신 3등급 (평균내신백분비4.5%차이)	내신 2등급-내신 3등급과 같은 효과를 내는 수능 점수 ()안은 200점 만점의 점수	내신 2등급-내신 3등급과 같은 효과를 내는 수능 점수 ()안은 300점 만점의 점수
40 : 30 : 30	1000점 중 점수	5	10.7(7.13)	7.70(7.70)
	회귀계수	0.065	0.041	0.038
45 : 25 : 30	1000점 중 점수	5.63	6.61(5.29)	6.87(6.87)
	회귀계수	0.058	0.0492	0.038
50 : 20 : 30	1000점 중 점수	6.25	3.80(3.80)	6.16(6.16)
	회귀계수	0.052	0.0615	0.038
50 : 25 : 25	1000점 중 점수	6.25	5.95(4.76)	4.28(5.13)
	회귀계수	0.052	0.0492	0.0456

〈표 13〉 내신 3등급과 4등급 차이가 CPA에 미치는 효과와 같은 효과를 미치는 수능 점수의 차이와 본고사 점수의 차이
(1000점 만점 중 차지하는 점수)

내신 : 수능 : 본고사의 비중		내신 3등급-내신 4등급 (평균 내신 백분비 5.5% 차이)	내신 3등급-내신 4등급과 같은 효과를 내는 수능 점수 ()안은 200점 만점의 점수	내신 3등급-내신 4등급과 같은 효과를 내는 수능 점수 ()안은 300점 만점의 점수
40 : 30 : 30	1000점 중 점수 회귀계수	5 0.065	13.08(8.72) 0.041	9.41(9.41) 0.038
45 : 25 : 30	1000점 중 점수 회귀계수	5.63 0.058	8.1(6.48) 0.0492	8.39(8.39) 0.038
50 : 20 : 30	1000점 중 점수 회귀계수	6.25 0.052	4.65(4.65) 0.0615	7.53(7.53) 0.038
50 : 25 : 25	1000점 중 점수 회귀계수	6.25 0.052	7.25(5.81) 0.0492	5.225(6.27) 0.0456

50 : 20 : 30 그리고 50 : 25 : 25로 상정하여 내신등급 1등급의 차이와 GPA에 같은 효과를 미치는 수능 성적의 차이, 본고사 성적의 차이를 계산하고, 이 3개의 값 차이가 입시총점 1000점 만점 중 차지하는 비중을 <표 11>~<표 13>로 계산하였다. <표 11>~<표 13>의 결과는 내신의 비중을 현행 40%에서 45% 정도로 상향 조정하는 것이 바람직하다는 것을 시사한다.

VI. 고교내신등급의 차등점수화의 필요성

고교내신 성적은 대학학업 성취도를 나타내는 GPA에 영향을 미치는 설명변수이다. 이상적인 상황에서 고교내신등급이 대학의 GPA의 예측지표라면 대학에 들어와서도 고교내신등급이 그대로 유지될 것이라고 생각할 수 있다. 즉, 이런 이상적인 상황에서는 내신 1등급 학생들이 입학당시 인문계열에서 28.4%를 차지할 경우 1년 후의 GPA분포에서도 상위 28.4%를 차지할 것이다. 이러한 이상적인 상황을 가정하여 고교내신등급과 유사한 대학내신등급을 정의하였고, 실제 상황이 이상적인 상황에서 얼마나 이탈되어 있는가를 관찰하고자 한다.

연세대학교에 입학한 학생들의 고교내신등급의 분포를 계열별로 살펴보면 <표 14>와 같다. <표 14>에서 조정된 누적 백분비는 인문계열에서 내신 4등급까지를, 자연계열에서는 내신 5등급까지를 전체로 간주할 때 조정된 누적 백분비이다. 인문계열에서는 내신 4등급까지가 전체의 88.8%를 차지하고, 자연계열에서는 내신 5등급까지가 전체의 90.4%를 차지하므로 이 학생들만을 대상으로 고교내신과 대학내신 간의 관계를 살펴보기로 하겠다.

<표 14> 1994학년도 입학생의 계열별 내신등급 백분비 표

인문계열의 내신등급(1~4) 백분비 표					자연계열의 내신등급(1~5) 백분비 표				
내신	빈도	백분비	누적 백분비	조정된 누적백분비(%)	내신	빈도	백분비	누적 백분비	조정된 누적백분비(%)
1	577	33.3	33.3	37.4	1	441	23.8	23.8	26.3
2	491	28.3	61.6	69.3	2	498	26.9	50.7	56.1
3	308	17.8	79.3	89.3	3	373	20.1	70.8	78.3
4	165	9.5	88.8	100	4	225	12.1	82.9	91.8
					5	138	7.4	90.4	100

1. 대학내신등급의 산정

대학내신등급을 산정하는 방법은 다음과 같다. i 고교에서 내신등급 j 를 받은 학생들의 집합을 S_{ij} 라 하자. S_{ij} 에 속한 학생들은 대학수학능력면에서 동질적이라고 가정한다. $C(A)$ 는 유한집합 A 를 구성하는 원소의 개수를 나타낸다고 하자. 그러면 다음과 같이 n_{ij} 와 D 를 정의한다.

$$n_{ij} = C(S_{ij})$$

$$D = \{(i, j), n_{ij} \geq 1\}$$

D 의 원소를 고교-내신등급으로 부르기로 한다.

D 에 속하는 (i, j) 에 대하여 S_{ij} 에 포함되는 학생들의 GPA를 $g_{ij1}, \dots, g_{ijn_{ij}}$ 라고 하자. 그러면 S_{ij} 군의 대학 GPA를 다음과 같이 정의되는 \bar{g}_{ij} 로 나타내기로 한다.

$$\bar{g}_{ij} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} g_{ijk}$$

그러면 D 에 속하는 모든 (i, j) 에 대하여 다음의 집합 G 를 구성할 수 있다.

$$G = \{\bar{g}_{ij}, (i, j) \in D\}$$

G 의 원소들을 크기순으로 나열하여 인문계열의 경우에는 <표 14>에 따라 상위 37.4%를 대학내신 1등급, 그 다음 상위 31.95%를 대학내신 2등급 등으로 정한다.

인문계열의 경우 D 에 속하는 (i, j) 에 대하여 i 는 $i = 1, \dots, 222$ 까지 값을 취하였고, $C(D) = 793$ 이었다. $C(D) < 222 \times 4$ 인 이유는 모든 고교가 학생들을 4개 등급 모두에 걸쳐서 연세대학교에 진학시키지는 않았기 때문이다. 인문계열과 자연계열 각각에서 대학내신 각 등급에 고교-내신등급의 숫자는 <표 15>와 같다.

<표 15> 계열별 대학내신등급별 고교-내신의 분포

인문계열(1,541명)				자연계열(1,657명)			
대학 내신	백분비(%)	누적 백분비(%)	고교-내신 등급수	대학 내신	백분비(%)	누적 백분비(%)	고교-내신 등급수
1	37.4	37.4	297	1	26.3	26.3	229
2	31.9	69.3	249	2	29.8	56.1	259
3	20.0	89.3	159	3	22.2	78.3	193
4	10.7	100	88	4	13.5	91.8	118
계			793	5	8.2	100	71
				계			870

2. 고교내신 백분비와 GPA의 상관분석

고교내신 백분비는 고교내신등급을 결정하는 席次累加百分比로서 상위 3%까지는 1등급, 3.01%~7.00%까지는 2등급 등으로 내신등급을 결정한다. 고교내신 백분비가 입시자료에 포함되어 있으므로, 고교내신과 대학내신 간의 상관계수를 구함에 있어서 굳이 離散化된 내신등급을 사용하지 않고 원자료인 고교내신 백분비와 대학 GPA 간의 상관계수를 구하기로 하겠다.

또한 고교내신 백분비와 대학 GPA 간의 상관계수를 고교별로 계산하여 고교별 상관계수의 분포를 살펴볼 수도 있다. 고교내신의 효과가 고교별로 큰 차이가 없다면 고교별 상관계수의 분포도 분산이 크지 않으리라는 것을 예측할 수 있다. 먼저 계열별로 고교내신 백분비와 GPA의 상관계수를 고교별로 구하여 그 기술 통계량을 살펴보면 <표 16>과 같다.

<표 17>을 보면 고교별 상관계수의 분포에서 분산이 매우 큰 것을 알 수 있고, 이는 고교내신의 효과가 일률적이 아님을 강력하게 증거하고 있다.

고교내신의 효과가 일률적이 아니라는 것은 또 다른 분석을 통하여서도 알 수 있

<표 16> 고교별 고교내신 백분비와 GPA의 상관계수의 기술통계량

	고교수 ¹⁾	평균	표준편차	최소값	제사분위수	중위수	제3사분위수	최대값
인문	106	-0.262	0.4708	-0.9926	-0.6284	-0.3806	-0.0076	0.9982
자연	219	-0.336	0.4630	-0.9990	-0.6699	-0.4102	-0.0443	0.9994

1) 입학인원 수가 2명 이하인 고교는 제외함.

<표 17> 계열별 고교내신 백분비와 GPA의 상관분석표

분석 대상		인문 계열	자연 계열
전 체	학생수	1735	1853
	상관계수	-0.17	-0.18
입학생 3명 이상 고교 대상	학생수(학교수)	860(106)	1588(219)
	평균	-0.282	-0.336
	표준편차	0.046	0.031
입학생 4명 이상 고교 대상	학생수(학교수)	809(89)	1390(153)
	평균	-0.340	-0.313
	표준편차	0.039	0.031
입학생 5명 이상 고교 대상	학생수(학교수)	765(78)	1278(125)
	평균	-0.318	-0.317
	표준편차	0.042	0.033

〈표 18〉 상관계수별 고교내신등급과 GPA원자료

상관계수 -0.9689			상관계수 -0.47450			상관계수 0.01140		
	고교내신등급	GPA		고교내신등급	GPA		고교내신등급	GPA
1	1	3.625	1	1	3.805	1	1	3.260
2	1	3.130	2	1	3.275	2	1	2.180
3	1	3.315	3	1	3.655	3	2	1.570
4	2	3.100	4	1	3.885	4	2	2.620
5	4	2.285	5	1	3.650	5	2	1.680
6	4	2.600	6	1	2.285	6	3	2.975
7	6	1.505	7	1	3.650			
			8	2	3.275			
			9	2	2.810			
			10	2	2.730			
			11	3	3.310			
			12	3	2.650			
			13	6	2.520			

다. 〈표 17〉은 고교별로 불력을 하지 않은 채로 고교내신 백분비와 GPA의 상관계수를 구하고, 이를 고교별 상관계수들의 평균과 비교하고 있다. 〈표 17〉에서 보듯이 고교별로 불력을 구성하지 않으면 인문계열에서 상관계수가 -0.17 이지만, 고교별로 불력을 구성하여 고교별 상관계수의 평균을 구하면 -0.282 부터 -0.340 까지로 상관계수의 절대값이 훨씬 커지는 것을 알 수 있다. 이는 고교별로 군집효과(cluster effect)가 있음을 증거하고 있다. 따라서 고교에 따라서 어느 고교는 고교내신 성적이 대학의 GPA와 높은 상관관계를 가지는 반면, 어느 고교는 전혀 상관관계를 가지지 않고, 심지어는 역상관관계를 나타내는 고교도 있다. 고교내신 백분비와 GPA의 상관계수가 높은 고교, 중간 정도인 고교, 거의 0에 가까운 고교를 예로 들어 원 자료를 살펴보면 〈표 18〉과 같다.

3. 고교내신등급과 대학내신등급의 결합도

VI. 1절에서 대학내신등급을 정의하고 이를 산정하는 방법을 기술한 바 있다. 이제 각 고교-내신등급 S_j 군에 속하는 학생들의 평균 GPA, \bar{g}_j 를 계산하여 대학내신등급을 할당하고, 대학내신등급과 고교내신의 결합도를 계산하여 보기로 한다.

〈표 19〉는 계열별로 고교내신등급과 대학내신등급의 분할표를 나타내고 있다. 각 고교-내신 등급에 해당하는 학생들이 2인 이상인 경우, 즉 $n_{ij} \geq 2$ 인 경우에만 분석에 포함시켰기 때문에 전체고교내신등급의 숫자가 $C(\{(i, j) ; n_{ij} \geq 1\})$ 의 경우보다는 작다. 〈표 19〉에서 인문계의 경우 H고교가 1등급 학생 5명, 2등급 학생 3명, 3등급 학생 4명을 연세대학교에 입학시켰다고 하자. 1등급 학생 5명의 평균 GPA는 대학내신등급으로 1등급에 해당하고, 2등급 학생 3명은 대학내신등급 3등급에

〈표 19〉 고교내신등급과 대학내신등급의 비교

인문계						자연계							
고교 내신등급	대학내신 등급					계	고교 내신등급	대학내신 등급					계
	1	2	3	4				1	2	3	4	5	
1	83					147	1	34					109
2		50				119	2		44				116
3			18			66	3			28			83
4				6		28	4				8		45
계						360	5					5	31
							계						384

해당하고, 3등급 학생 4명은 대학내신으로 4등급에 해당한다고 하면 H고교는 (1, 1), (2, 3), (3, 4) 칸에 각각 1의 dot수를 갖게 된다. 〈표 19〉를 보면 자연계의 고교내신등급 4등급, 5등급을 제외하고는 모두 대각항의 dot수가 제일 많은 것을 알 수 있으나, 비대각항의 dot수도 상당히 큰 것을 알 수 있다. 순위형 분할표에서 결합도를 나타내는 쏘머스의 D를 계산하면 〈표 20〉과 같다.

쏘머스의 D값은 상당히 완만한 결합도를 나타내 주고 있다. 〈표 19〉에서 각 칸에 해당되는 고교의 명단을 파악하여 일반적으로 대각선 왼쪽 아랫 부분에 위치하는 고교와 대각선 오른쪽 윗부분에 위치하는 고교에 대하여서는 고교내신 성적에 대하여 차등점수화를 고려하는 것이 타당하다.

〈표 20〉 계열별 고교내신등급과 대학내신등급의 결합도

	쏘머스 D (표준오차)
인문계열	0.303 (0.044)
자연계열	0.172 (0.042)

VII. 연구의 한계 및 추후 연구제안

본 연구는 1994년도에 부활된 대학별 본고사와 수학능력시험, 그리고 특별전형 제도에 대한 여러 가지 가설들을 검증하고 있고, 1994학년도 대학입학생의 대학수학능력을 1년간의 GPA로 측정하였다. 대학수학능력을 첫 1년간의 GPA로 측정하기에는 대학 1학년의 고유한 학업 분위기상 다소 미흡한 점이 있다고 판단한다. 이 부분에 대하여서는 누적된 자료를 통하여 지속적인 연구를 수행해 보는 것이 바람

직하다.

아울러 *GPA*에 영향을 미치는 개개인의 미시적 자료, 즉 학부모의 경제적 여건, 학생의 동아리 활동 그리고 그 외에 시간 관리를 측정할 수 있는 제 변수들에 대한 자료가 전혀 없으므로 *GPA*를 반응변수로 한 회귀모형은 적합도에 한계가 있었다 ($R^2 = 0.250$). 이와 관련하여 장기적인 안목에서 입학생의 일정비율을 표본조사하여 4년간의 대학생활을 추적조사하는 패널자료의 구성이 요청된다. 또한 고교내신의 차등점수화의 필요성에 대하여서는 통계적으로 정교한 임의계수회귀⁵⁾(Random Coefficient Regression) 모형을 구성하여 임의효과에 대한 검정을 실시할 수 있을 것이다.

◆ 參 考 文 獻 ◆

1. Anderson, T. W., *An Introduction to Multivariate Statistical Analysis*, Wiley. N. Y., 1958.
2. Brown, M. B. and Forsythe, A. B., "Robust test for the equality of variance," *Journal of the American Statistical Association*, Vol. 69, 1974.
3. Neter, J. and Wasserman, W. and Kutner, M. H., *Applied Linear Regression Models*, Irwin, Homewood, 1989.
4. Longford, Nicholas, T., *Random Coefficient Models*, Oxford University Press Clarendon Press, Oxford, 1993.
5. SAS Institute Inc, *SAS/STAT User's Guide*, Version 6, Fourth Edition, Vol. 1, Cary, 1990.
6. SAS Institute Inc, *SAS Procedures Guide*, Version 6, Third Edition, Cary, 1990.

5) Longford [4] 참조.