韓國數學敎育學會誌 시리즈 A <數學敎育>
1994. 12. 제 33권, 제 2호, 167-175.

Journal of the Korea Society of Mathematical Education
Series A: Dec. 1994. Vol. 33, No. 2, 167-175.

결측치를 가진 목표지향형 평가모델에서 수학학습능력의 평가에 관한 연구

홍 석 강 (동국대학교)

<요약> 결측치를 가진 회귀모형의 모수 추정법을 이용하여 목표지향형 평가 모델에서 기초고사(X)와 신고사(Ⅰ)(Y), 신고사(Ⅱ)(Z)등 두 개 이상의 고사지로 이루어진 고사집에서 기초고사에는 결측치가 없고 신고사(Ⅰ), 신고사(Ⅱ)등에는 결측치가 있는 경우 모수의 최우추정량 계산법을 논하고 E.M.알고리즘과 평가치의 회귀방정식화에 의하여 우리나라 중등학교 학생의 수학학습능력과 수학적사고력의 크기를 변별하며 학생들의 진능력이 반영된 평가모델과 최종 성적을 평가 할 수 있는 계산법을 제시하였다.

Ⅰ. 서론

목표지향평가 (crierion referenced evaluation)는 학습평가의 결과를 해석함에 있어서 사전에 구체화된 목표에 비추어 해석하는 평가방법으로써 그 목표는 학습의 계획과정에서 설정된 학습목표일 수도 있고 수행규준 일 수도 있는데 이 평가는 오늘날 개인적으로 널리 큰 비판없이 시행되고 있는 평가법, 즉 개인차의 변별, 상대적 비교, 객관성등을 지나치게 강조함으로써 학습자들의 외현적 동기 유발에만 유의한 결과 인간의 발전성에 대한 신념이나 교육의 힘에 대한 신념을 흐리게 하고 교수-학습의 개선 기능을 약화시키는 단점을 가진 규준지향 평가법(norm referenced evaluation)보다 인성과 지적개발에 많은 장점을 가진 평가법인데 이 평가법에는 다음과 같은 특징이 있다. 이 평가에서는 (1) 참다운 학력평가가 가능하며 (2) 교육의 힘에 대한 신념을 바탕으로 하고 있으며 (3) 현대적 학습이론에 부합하는 점, 즉 외현적 동기유발보다는 동기적 유발 방법을 강조하고 (4) 교수-학습의 개선에 도움이 된다는 점 등을 들 수 있다. 그러나 이와 같이 목표지향평가법의 좋은 특징도 많으나 무엇보다도 어려운 점은 목표지향의 절대기준 설정이 그리 쉽지 않다는 점, 즉 절대기준의 설정을 위하여는 각 교과의 학습구조에 대한 분석과 학습과제의 위계화 분석이 필수적으로 요구되므로 그 작업이 고도의 전문성을 필요로 하며 각교과의 학문구조나 학습 과제에 대한 분석자료는 교육과정이나 교과서에 제시되어 있지만 이를 분석하고 체계화하는 과정에서 필연적으로 교사의 교육관이 작용하게 되므로 결과적으로 분석의 결과가 서로 다르게 평가 될 수 있는 점이다. 그러므로 이러한 결점을 보완하기 위하여 각 교과 분야에서 그 교과마다 특징을 살린 교육목표 이원분류표, 교과내용 분류표 행동목표의 분류표를 작성하고 평가도구의 구비조건으로써 문항지의 내재적 준거(intrinsic criterion)인 타당도, 신뢰도, 양호도, 객관도 등이 반영된 양호한 검사 문항을 가진 평가모델을 제작하며 특히 수학교육 현장에 이 모델을 적용하여 수학학습능력과 사고력을 증진시키는 효과적인 지도법을 제시하고 학생들의 진능력이 반영된 평가법을 고안하여 지정된 최종학습목표에 성공적으로 도달할 수 있도록 하여야 할 것이다.

다음 목표지향평가법에 관한 여러 연구 분야를 고찰하면 (1) 평가지의 내재적 준거면에서 본 규준지향평가나 목표지향평가의 고사형 및 평가치의 측정에 관한 비교 연구는 Block (1971), Carver(1974), Ebel(1971), Nitko(1974), Hambleton et. al(1978) 등이 평가의 장단점을

비교한 결과를 제시한 바 있고, (2) 외재적 준거면에서 본, 즉 출제 방법이나 평가도구의 모델에 의하여 나타나는 효과를 더욱 중요시하여 수험자의 학습능력을 최종 목표 수준까지 도달시키기 위한 평가모델의 개발에 관한 연구는 Alkin(1994), Milman(1973) 등이 새로운 문항 배치법 개발과 그 모델을 심리측정 분야에 응용한 예를 들고 있으며 또 우리나라 중등학교 수학교과에서 수학적 사고력문제 중 이해력의 심도 크기를 측정하기 위한 고사모형에 관하여는 홍석강(1988, 1990) 등의 연구결과를 들 수 있고, (3) 결측치를 가진 목표지향형 배치 모델의 최우추정량을 계산하는 연구는 Rubin(1976), Thyer(1982) 등이 기초고사에는 결측치가 없고 신고사에만 결측치가 있는 경우 모수를 추정하는 법을 제시하였는데, 즉 이 경우는 신고사에 결측치가 있으므로 평가치가 모두 관측되어진 완비형 모델의 경우처럼 모수를 직접적으로 추정하기 어렵기 때문에 그 불완비형 배치 모형을 보다 작은 부분 고사형으로 분해(factoring)하면서 그 소불완비형 평가 모델의 우도(likelihood)로 분해하고 그 우도의 모든 곱(product of likelihood)으로써 소불완비형 모델을 모두 합하여 표준완비형 모델화시켜 최우 추정량을 계산함을 보이고 있다. 다음 (4) 평가치 회귀방정식화(test equating)의 연구에는 Thyer(1982), Dempster et. al. (1977)등이 각 부분고사의 결합 공분산 행렬의 최우추정치를 계산하는 모수추정법을 제시하고 그 평가치를 선형 방정식화(linear equating)한 후 구한 회귀방정식의 회귀계수들의 안정성(stability)과 예민성(sensitivity)의 크기에 관하여 논하고 있다. 그러므로 이 논문에서는 위의 여러 학자들이 연구한 목표지향형 평가모델의 모수추정법을 참고하여 그 추정법을 우리나라 중등학교 학생들의 수학 학습능력이나 수학적 사고력을 변별할 수 있는 평가모델에 적용하는 법을 제시하며 학생들의 문제 해결력에서 더욱 어려운 문제를 해결한 학생들에게 그들의 진능력(true ability)이 반영된 공정한 평가를 주어 학생들의 학습능력을 증신시키고 나아가서는 지정된 최종 학습목표 수준까지 효과적으로 도달시키게 함을 연구의 목적으로 한다.

Ⅱ. 연구의 내용

A. 기초고사 때는 결측치가 없고 신고사에만 결측치가 있는 목표지향형 평가 모델과 그 모수 추정법

Rubin(1976), Thyer(1982) 등은 각 문항 부표본들의 표본의 크기 n_1, n_2, n_3, …인 기초고사와 신고사로 구성된 고사집의 평가 문항에서 기초고사에는 결측치가 없고 신고사에만 결측치가 있는 목표지향형 평가모형과 그 평가치의 평균 벡터, 공분산 행렬, 편상관계수등 모수의 최우추정량을 계산하는 과정을 다음과 같이 제시하였다.

(1) 관측치와 결측치에 대한 가정 :

(가) 관측치와 결측치는 항상 무작위로 일어난다.

(나) 동일한 관측치는 평가치의 어떤 다른 변량과도 서로 무상관이다.

(2) 평가치의 변량 벡터 $(X_1,\cdots, X_s, Y_1,\cdots, Y_p, Z_1, \cdots, Z_q)$를 $(X_1, \cdots, X_s, Y_1, Y_p \cdots,)$와 $(X_1, \cdots, X_s, Z_1, \cdots, Z_q)$로써 그 변량들의 평균 벡터와 공분산 행렬로 각각 분해하고 X상에서 Y와 Z의 회귀 계수의 계산을 위하여 부표본별로 본 Y나 Z의 평가치의 열관측치 벡터 중 적어도 한 개의 열 이상이 있을 것.

(3) 추정할 수 없는 모수들의 분리및 계산

다변량 결합분포를 갖는 평가치 변량들 X_i, Y_i, Z_i $(i=1, 2, \cdots)$ 중

(가) 변량 X는 관측되어졌으므로 평균 벡터 $\widehat{\mu_x}$, 공분산 행렬 $\underset{\sim}{\sum} xx$, 회귀계수 $\widehat{b_i}$ 등 X의 최우추정치는 쉽게 계산되어 질 수 있다.

(나) 관측되어진 변량 X가 주어진 조건부 변량 Y와 Z간의 상관 정도의 크기를 측정하고져

Y, Z와 X와의 결합 분표로써 결측치 때문에 직접 계산하기 어려운 모든 모수를 추정하고져 할 때 X가 조건부로 주어진 것을 비조건부화 (unconditioning)화 하는 과정을 거친후 추정가능화 시켜 그 계산된 모수를 계산하고 그 모수들의 크기로써 범위를 지정하여 가장 적당한 값을 선택하여 변량 X와 Y, Z등의 상관의 정도를 관찰한다. 즉 한 예로써 (다)의 계산과정 결과 X가 비조건부화 되어진 값으로써 (X, Y)의 공분산 행렬 $\sum\limits_{\sim} xy$, X와 $(Y_1+Y_2+Y_3)$의 다중상관 행렬 등을 계산할 수 있는 데 이 중상관계수나 편상관계수에 대한 가능한 값들의 범위를 구한 후, 그 값 가운데 적당한 값을 선택 대채하여 다른 모수의 추정을 위한 통계적 계산을 할 수 있다. 또 이 값들의 범위를 실제로 지정할 수 있는 계산근거인 Y_i와 Y_j의 상관관계는 그 사전 확률(priori)의 크기에 따라 좌우되므로 Y_i와 Y_j의 상관관계를 더욱 제한되어진 값으로 한정시키므로써 그 한정된 범위내에서 유용한 값을 선택하면 더욱 효과적이 될 것이다. 최종 평가 결과로써 신고사 관측치의 합인 $(Y_1+Y_2+Y_3+\cdots)$로 학습자의 최종평가치가 어느정도인가를 변별할 수 있으므로 지정된 최종 학습목표 수준에 도달 여부를 판정할 수 있겠다.

B. 주 연구 내용

이 절에서는 앞 절에서 제시한 Rubin(1976)와 Thyer(1982) 등의 목표지향형 평가 모형과 그 모수 추정법을 이용하고 Dempster et. al.(1977)가 제시한 E.M. 알고리즘과 Sweeping Operation을 이용한 평가치의 회귀방정식화 법으로써 우리나라 중등학교 학생들의 수학학습능력 변별화 및 목표지향형 평가에서 목표수준까지의 학습능력판정을 위한 평가의 통계적 해석법을 제시하고자 한다.

[정의] E.M. 알고리즘

두 개의 표본공간 X와 Y에서 다대일 사상인 X → Y가 성립할 때 불완비형 자료 (incomplete data)가 생긴다. 여기서 Y는 실현가능치이고 X는 모두를 포함하는 경우도 있으나 X → Y(X)의 가정하에서 완비형 자료 X{x∈X(t), y = Y(X)}는 Y에 의해서만이 구해질 수 있음을 가정한다. 이 때,

$$g(Y/\ \phi)=\int_{X(Y)} f(X/\ \phi)dx, \quad \cdots\cdots\cdots (1)$$

(단 f(X/ ϕ)는 모수 ϕ에 종속되는 표본밀도 X(Y)의 함수족(A family of sampling densities depending on parameter ϕ))

<표 1> Rubin과 Thyer의 모델

문항의 부표본 (Sub-samples)	표본의 크기 (Sample Size)	기초고사 (Old Test) $X_1, \cdots, X_p$	신고사 (New Test) Y_1, Y_2, Y_3
1	1 ⋮ n_1	v ⋯ v ⋮ v ⋯ v	1 - - ⋮ 1 - -
2	n_1+1 ⋮ n_1+n_2	v ⋯ v ⋮ v ⋯ v	- 1 - ⋮ - 1 -
3	n_1+n_2+1 ⋮ $n_1+n_2+n_3$	v ⋯ v ⋮ v ⋯ v	- - 1 ⋮ - - 1
	⋮	⋮	⋮

v는 관측되어진 문항, -는 관측되어지지 않은 문항

<표 2>

$$\underset{\sim}{D} = \begin{array}{c|c|ccc} & \underset{\sim}{X} & \underset{\sim}{b_1} & \underset{\sim}{b_2} & \underset{\sim}{b_3} \\ \hline & \underset{\sim}{b_1}^T & V_1 & ? & ? \\ \underset{\sim}{Y} & \underset{\sim}{b_2}^T & ? & V_1 & ? \\ & \underset{\sim}{b_3}^T & ? & ? & V_1 \end{array}$$

X가 Swept out 될 (Y, X)의 교적행렬

<표 3> Thyer의 모델

문항의 부표본 (Subsample)	표본의 크기 (Sample Size)	기초고사 (old Test) $X_1 \cdots X_s$	신고사(Ⅰ) (New Test(Ⅰ)) $Y_1 \cdots Y_p$	신고사(Ⅱ) (New Test(Ⅱ)) $Z_1 \cdots Z_q$
1	n_1	v · · · v	v v · · · –	– – · · · –
2	n_2	v · · · v	– v · · · v	– – · · · –
3	n_3	v · · · v	v – · · · v	– – · · · –
4	n_4	v · · · v	– – · · · –	v v · · · –
5	n_5	v · · · v	– – · · · –	– · · · v v
6	n_6	v · · · v	– – · · · –	v · · · – v
·	·	·	·	·
·	·	·	·	·
·	·	·	·	·

v는 관측되어진 문항, -는 관측되어지지 않은 문항

<표 4> 평균치, 분산, 공분산행렬의 최우추정치를 나타내는 분할행렬

	1	$\underset{\sim}{X}^T$	$\underset{\sim}{Y}^T$	$\underset{\sim}{Z}^T$
1	1	$\widehat{\underset{\sim}{\mu}_X}^T$	$\widehat{\underset{\sim}{\mu}_Y}^T$	$\widehat{\underset{\sim}{\mu}_Z}^T$
$\underset{\sim}{X}$	$\widehat{\underset{\sim}{\mu}_X}$	$\hat{\underset{\sim}{\Sigma}}_{XX}$	$\hat{\underset{\sim}{\Sigma}}_{XY}$	$\hat{\underset{\sim}{\Sigma}}_{XZ}$
$\underset{\sim}{Y}$	$\widehat{\underset{\sim}{\mu}_Y}$	$\hat{\underset{\sim}{\Sigma}}_{YX}$	$\hat{\underset{\sim}{\Sigma}}_{YY}$	?
$\underset{\sim}{Z}$	$\widehat{\underset{\sim}{\mu}_Z}$	$\hat{\underset{\sim}{\Sigma}}_{ZX}$	?	$\hat{\underset{\sim}{\Sigma}}_{ZZ}$

?는 Y와 Z의 미지의 공분산 행렬을 나타냄

<표 5> $(X_1,\cdots,X_s)$상에서 $(X_1,\cdots,X_s, Y_1,\cdots, Y_p, Z_1,\cdots,Z_q)$ 의 교적행렬

	$\underset{\sim}{X}^T$	$\underset{\sim}{Y}^T$	$\underset{\sim}{Z}^T$
$\underset{\sim}{X}$	$\hat{S}_{XX}^{-1}$	$\hat{\delta}_{Y/X}$	$\hat{\delta}_{Z/X}$
$\underset{\sim}{Y}$	$-\hat{\delta}_{Y/X}^{T}$	$\hat{\underset{\sim}{\Sigma}}_{Y/X}$	?
$\underset{\sim}{Z}$	$-\hat{\delta}_{Z/X}$	?	$\hat{\underset{\sim}{\Sigma}}_{Z/X}$

?는 Y와 Z 변량에 편표준편차의 최우추정량을 곱한 것으로써 추정이 불가능한(inestimable) 편상관계수를 나타냄

를 정의하면 E.M. 알고리즘이란 g(Y/ ϕ)를 최대화(maximization)한 ϕ의 값을 구하는 것에 있다.

[정리1] 정규지수함수족
(Regular Exponential Family Form)

$$f(X/\phi) = b(X)\exp[\ \{\phi(t(X))^T\}/a(\phi)] \ \cdots (2)$$

(단, ϕ는 크기, 1 x r인 모수벡터, t(X)는 크기 1 x r인 완비형 자료의 충분통계량(complete data sufficiet statistics), T는 전치행렬)

에서

(1) 기대치 과정 (expectation step)

$$t^{(P)} = E\{t(X)\ /\ Y,\ \phi^{(P)}\} \ \cdots\cdots (3)$$

를 구하여 t(X)를 추정하고

(2) 최대화 과정 $E\{t(X)\ /\ \phi\}$ …… (4)

의 방정식의 해로써 $\phi^{(P+1)}$를 구한다.

여기서 f(X/ ϕ)가 $\phi \subset \Omega$에 대하여 밀도함수를 정의하는 r차원의 볼록집합(convex set)인 Ω상에서만 정의 될 때 이 f(X/ ϕ)는 정규형(regular form)이라 한다.

(증명)

식 (1)과 (2)에 의하여

$$L(\phi) = \log g(Y/\phi)$$

를 최대화시키는 ϕ의 추정치 ϕ^*를 구하기 위하여 지수함수족을 k로 표시하여

$$k(X/Y,\ \phi) = f(x/\phi)/g(Y/\phi)$$

일 때

$$L(\phi) = \log f(X/\phi) - \log k(X/Y,\ \phi)$$

이다. 그러면 지수함수형 $k(X/Y,\ g)$는 식 (2)에서

$$k(X/Y,\ \phi) = b(x)\exp[\ \{\phi(t(x))^T\}/a(\phi/Y)]$$

$$(\text{단},\ a(\phi/Y) = \int_{X(Y)} b(x)\exp[\ \phi\{t(X)\}^T]\ dx)$$

에서 $f(X/\phi)$와 $k(X/Y,\ \phi)$는 모두 동일한 모수 ϕ와 충분통계량 $t(X)$를 가지므로

$$L(\phi) = -\log a(\phi) + \log a(\phi/Y) \cdots\cdots\cdots (5)$$

$$a(\phi) = \int_X b(X)\exp[\ \phi\{t(X)\}^T]\ dx \cdots\cdots (6)$$

그러므로 식 (5)와 (6)을 미분하여

$$D\ \log a(\phi) = \frac{\partial}{\partial\phi}\log a(\phi) = E(t/\ \phi)$$

$$D\ \log a(\phi/Y) = E(t/Y,\ \phi)$$

이고 식 (5)에 대입하면

$$D\ L(\phi) = -E(t/\phi) + E(t/Y,\ \phi) \quad \cdots\cdots (6')$$

이다.

식 (6′)을 E.M. 알고리즘 정의에 의하여 반복적으로 계산하여 $\phi^{(P)} = \phi^{(P+1)} = \phi^*$인 극한치 즉 $E(t/\phi^*) = E(t/Y,\ \phi^*)\ \phi^*$를 얻고 이것은 $D\ L(\phi)=0$을 만족한다.

[정리 2] Sweeping 알고리즘

(1) Sweep Operator : 행렬 $M_{P\times P}$에 k행 $(i\neq k)$과 k열$(j\neq K)$에 대하여 다음 계산과정

$$n_{kk} = -\frac{1}{m_{kk}},\quad n_{ik} = \frac{m_{ik}}{m_{kk}},\quad n_{kj} = \frac{m_{kj}}{m_{kk}}$$

$$n_{ij} = m_{ij} - \frac{m_{ik}\cdot m_{kj}}{m_{kk}}$$

을 행하여 얻은 새로운 행렬 $N_{P\times P}$는 $Swp[\ k]\ M$으로 표현되고 $M_{P\times P}$에 $k_1,\ k_2,\ \cdots, k_t$의 연산을 연속적으로 행하면 $Swp[\ k_1,\ k_2,\ \cdots,\ k_t]\ M$을 얻는다.

(2) Reverse Sweep Operator : 행렬 $N_{P\times P}$에 k행 $(i\neq k)$과 k열$(j\neq K)$에 대하여 다음의 역 Sweep Operation의 계산 과정

$$m_{kk} = -\frac{1}{n_{kk}},\quad m_{ik} = -\frac{n_{ik}}{n_{kk}},\quad m_{kj} = -\frac{n_{kj}}{n_{kk}}$$

$$m_{ij} = n_{ij} - \frac{n_{ik}\cdot n_{kj}}{n_{kk}}$$

을 행하여 얻은 행렬 $M_{P\times P}$는 $RSwp[\ k]\ N$으로 표현되고 $N_{P\times P}$에 $k_1,\ k_2,\ \cdots,\ k_t$의 연산을 연속적으로 행하면 $RSwp[\ k_1,\ k_2,\ \cdots,\ k_t]\ N$을 얻는다.

지금 정의 1 의 E.M. 알고리즘을 이용한 회귀방정식을 표기 하면 기초고사(X), 신고사(Ⅰ)(Y), 신고사(Ⅱ)(Z)등의 변량

$X = \sum_{i=1}^{p}\sum_{j=1}^{n_i} X_{ij}$, 단, $j = n_1,\ n_2,\ \cdots,\ n_n$의 표본 으로 표기함

$Y = \sum_{i=1}^{n_i} Y_i$, 단, $i = n_1,\ n_2,\ \cdots,\ n_n$의 표본으로 표기함

$Z = \sum_{i=1}^{q_i} Z_i$, 단, $i = q_1,\ q_2,\ \cdots,\ q_n$의 표본으로 표기함

들로써 최종평가치의 합계인 T에 관한 모수를 계산하는 식은 다음과 같다.

$$T = Y + Z$$

$$\mu_{T_i} = \sum_{i=1}^{n_i}\mu_{Y_i} + \sum_{j=1}^{q_i} Z_i \cdots\cdots\cdots (7)$$

$$\sigma^2_{T_i} = \sum_{i=1}^{n_i} \sigma^2_{Y_i} + 2\sum_{i<j} cov(Y_i,\ Y_j) + \sum_{i=1}^{q_i} \sigma^2_{z_i}$$

$$+2\sum_{i<j} cov(Z_i,\ Z_j) + 2\sum_{i=1}^{n_i}\sum_{j=1}^{q_i} cov(Y_i,\ Z_j) \cdots (8)$$

또 한고사지내에서 기초고사(X) 평가치와 신고사(Y, Z) 평가치들로써 응시자의 수학적사고력 변별및 문제해결의 진능력을 검정하기 위한 평가치의 회귀방정식과 X와 Y_i의 중상관계수에 대한 최우추정량의 식은 다음과 같다.

$$Y^* = e_x(Y) = \mu_x + \frac{\sigma_X}{\sigma_Y}(Y-\mu_y) \cdots\cdots\cdots (9)$$

$$Z^* = e_x(Z) = \mu_y + \frac{\sigma_y}{\sigma_Z}(Z-\mu_z) \cdots\cdots\cdots (10)$$

$$\rho_{Y_i \cdot X} = \left[\ 1 - \frac{V_i}{\sigma^2_{Y_i}}\right]^{\frac{1}{2}} \cdots\cdots\cdots\cdots\cdots (11)$$

여기서 위의 식들을 이용하여 신고사(Y, Z)를 수학학습능력 또는 수학적 사고력 변별의 유형별 변수로 두고 각 평가치 변량들의 상관계수와 중상관계수들을 계산하여 변량 X와 Y, Z간의 상관의 정도를 측정한 후 응시자들의 진능력이 반영된 회귀방정식 추정치 Y^*와 Z^*를 계산 함으로써 학습목표들의 최종 수준까지의 도달여부를 판정할 수 있다.

다음예는 어느 대학 입시 평가 연구소에서 시행한 수학 II의 전국 정답율을 계산한것으로써 그 정답율의 난이도로써 그들의 수학문제 해결능력을 변별하기 위하여 그 평가치 X_1(상수), X_2(1회 고사), X_3(2회 고사), Y_1(정답율 42% 이상의 문항평가치), Y_2(정답율 42%미만의 문항평가치)로 1회고사와 2회고사에서 학력수준이 고득점인 문항평가치와 저득점인 문항평가치간의 중상관계수, 상관계수, 기초고사(X)에서의 문제 해결능력이 반영된 최종평가치 Y_x^1, Y_2^* 등을 계산 하였다.

[예] X_1 = 상수, X_2 = 수학(II)의 1회고사 문항평가치, X_3 = 수학(II)의 2회고사 문항 평가치, Y_1 = 정답율 42% 이상의 문항 평가치, Y_2 = 정답율 42% 미만의 문항 평가치

평균화 시킨 X의 교적 행렬은

$$C_{xx} = \begin{bmatrix} 1 & 39.4642 & 48.56 \\ 39.4642 & 1863.12 & 1973.45 \\ 48.56 & 1973.45 & 2586.09 \end{bmatrix}$$

이고 기초고사 X상에서 Y_1과 Y_2의 중회귀계수와 잔차 분산은 각각

$\underset{\sim}{b_1}^T = (19.8012, 0.2976, -0.0182)$, $V_1 = (77.54652)$

$\underset{\sim}{b_2}^T = (49.5714,\ 0.1701,\ 0.0285)$, $V_2 = (100.30403)$

이였다. 이 자료로써 추정 불가능한 공분산의 값을 계산하기 위하여 <표 2>의 ?대신 0을 대입한 후 편상관계수에 좌우되지 않는 최우추정치들을 계산하기 위해서 변량 X_2와 X_3에 대하여 정리(2)의 Reverse Sweep operation을 이행하여 다음의 행렬 E를 얻었다.

$$E = \begin{array}{c} \begin{array}{ccccc} X_1, & X_2, & X_3, & Y_1, & Y_2 \end{array} \\ \begin{bmatrix} -1 & 39.4642 & 48.56 & 30.661954 & 57.66822 \\ 39.4642 & 305.69692 & 57.068448 & 89.936757 & 53.625497 \\ 48.56 & 57.068448 & 228.0164 & 12.833672 & 16.20581 \\ 30.661954 & 89.936757 & 12.833672 & 104.07901 & 15.664002 \\ 57.66822 & 53.625497 & 16.20581 & 15.664002 & 109.88776 \end{bmatrix} \end{array}$$

이 행렬 E의 첫행에서 $(X_1, X_2, X_3, Y_1, Y_2)$는 평균 벡터의 최우추정치이며 대각 원소는 그 평가치 변량의 분산의 최우 추정치이다. 따라서 식 (11)에 의해 Y_i와 X의 중상관계수는

$$\rho_{Y_1 \cdot X} = 0.5049024\ ,\quad \rho_{Y_2 \cdot X} = 0.2953199$$

이고 각 평가치 변량간의 상관계수는

$\rho_{X_2 \cdot X_3} = 0.2161561$

$\rho_{X_2 \cdot Y_1} = 0.504209$, $\rho_{X_2 \cdot Y_2} = 0.8218203$

$\rho_{X_3 \cdot Y_1} = 0.0833079$, $\rho_{X_3 \cdot Y_2} = 0.1023795$

이다. 다음에는 식 (9)에 의하여 기초고사 X의 기초학력평가치가 반영된 최종평가치

Y_1^{*} = (17.6988, 28.2825, 30.2760, 12.6330, 49.3677, 20.3088, 23.9360, 50.2925, 51.0323, 31.7967, 48.3402, 44.5280, 41.8872, 45.7919, 41.6714, 44.6513, 49.8507, 41.1063, 39.2772, 26.5460, 42.3085, 55.2761, 32.8962, 40.4486, 48.9875, 42.3496, 36.6775, 49.2033, 39.7294, 50.7960, 43.4491, 43.2230, 37.6537)

Y_2^{*} = (40.7122, 32.8043, 39.8092, 38.6668, 48.3465, 35.8148, 43.2042, 55.8799, 72.6609, 31.8104, 71.0764, 59.6395, 43.6507, 77.9329, 30.4708, 27.6043, 37.6873, 26.0487, 36.7799, 44.1549, 42.5560, 26.8553, 80.1367, 47.2518, 54.3675, 51.0401, 57.5076, 43.2618, 55.4046, 74.2454, 56.2689, 61.0079, 57.7381)

를 얻을 수 있고 이 최종평가치들은 문항의 정답율로 본 응시자들의 최종목표수준의 학습능력을 측도 할 수 있는 평가치로 간주된다. 그리고 이 자료에서 관찰된 상관계수로써 $\rho_{X_2 \cdot Y_2}$가 가장 큰 것으로 보아 1회 기초고사 (X_2)에서 가장 어려운 문항들이 많이 출제된 것으로 추정되어지고 $\rho_{X_3 \cdot Y_1}$이 가장작은 값인 것으로 보아 2회 고사 (X_3)에서는 어려운 문항들이 별로 출제되지 않은 것으로 추정할 수 있으며 두 기초고사(X_2, X_3)사이에는 상관계수 $\rho_{X_2 \cdot X_3}$의 값이 작은 것으로 보아 두고사는 상호 상관이 없는 2회 시행의 동형고사를 치른것으로 생각되어진다. 뿐만 아니라 평가치 Y_1과 Y_2를 재평가하기 위하여 적합시킨 선형 회귀 방정식치 Y_1^{*}과 Y_2^{*}에서 Y_1^{*}가 Y_1보다 조금 높게 평가되고 Y_2^{*}는 Y_2보다 조금 낮게 평가된 것으로 보아 2회의 기초고사 (X_2, X_3)에서 재평가된 Y_1^{*}와 Y_2^{*}는 응시자들의 최종 문항 정답율의 산정 평가치로써 그들의 진능력을 평가할 수 있는 값이 됨을 알 수 있다.

Ⅲ. 결론

본고에서는 목표지향형 평가의 특징, 문항 배치법 및 여러 학자들이 개발한 평가치의 효과적인 해석법을 논하고 특히 Rubin(1976)와 Thyer (1982) 등이 개발한 모형과 모수의 최우추정량 계산법을 목표지향형 평가모델에 적용하여 우리나라 중등학교 수학교과에서 고3 응시자들의 수학 문제해결 능력을 변별 할수 있는 모델에 응용하였으며 그들의 학습능력이 최종목표 수준까지의 도달 여부를 판정하고 또 그들의 기초고사 성적이 반영된 진능력의 최종평가치를 선정하는 통계적 해석법을 제시하였는데 이 논문의 주요 연구결과와 기대효과 및 활용 방안을 요약하면 다음과 같다.

(1) 목표지향형 평가의 효과적인 운용은 오늘날의 교육적 관심대상인 영재 교육분야에서 월반, 영역별 속진, 각 교과내용의 심화학습등, 수월성 추구를 위한 특수재능교육 프로그램을 위한 평가도구와 검사도구등을 개발하는데 기여할 수 있다.

(2) Ⅱ장과 A, B절에서 논한 바와 같이 기초고사에는 결측치가 없고, 신고사에는 결측치가 있는 Rubin(1976)와 Thyer(1982)의 평가모델을 목표지향형 평가모델에 적용하고 정리 1, 정리 2 등을 이용하여 모수의 최우 추정량 계산을 더욱 용이하게 계산하는 해석법을 제시하였다.

(3) 모수 최우추정량계산 결과인 <표 2>를 얻기 위한 계산과정에는 회귀해석에 널리 이용되고 있는 결측치를 가진 회귀모형의 분산분석법을 결측치가 있는 목표지향형 평가모델의 경우에 응용한 것으로써 Sweeping 알고리즘을 적용하여 보다 직접적이고 간편하게 모형균벡터의 추정치와 공분산 행렬로 구성된 평가치 행렬등을 구할수 있는 연산과정을 도입하였다.

(4) (3)의 결과로써 결측치가 있는 경우 평균벡터, 중상관 계수등 추정이 어려운 모수들을 추정한 후, 그 모수들로써 평가치 (Y_1, Y_2)등을 선형 회귀화시켜 재평가함으로써 기초고사에서의 응시자들의 학습능력이 반영되고 또 학습목표의 최종학습 수준까지의 달성여부를 판정할수 있는 최종평가치 (Y_1^*, Y_2^*)를 산정하는 해석법을 제시하였다.

참 고 문 헌

홍석강(1988). 중등학교의 수학적사고력 증진을 위한 효과적인 수학지도법의시안. 한국수학교육학회지 수학교육, 제 26권, 제 2호, 15-41.

홍석강(1988). 수학교육에서 이해력 심도의 측정과 방법. 교육문제연구(동국대학교 교육문제연구소), 제 5집, 83-95.

홍석강(1990). 신구 두고사 평가치 변환에 의한 진분포와 모수추정에 관한 연구. 한국수학교육학회지 수학교육, 제 29권, 제 2호, 79-93.

홍석강 외 4인(1993). 중학교 수학교과의 수업모형, 수업방법, 평가방법및 평가도구 개발에 관한 연구 (연구보고 RR-92-Ⅱ-3). 한국교원대학교 부설 교과교육공동연구소.

Alkin, M.C. (1974). Criterion-referenced measurement. In C.W. Harris, M.C. Alkin & W.J. Popham (Eds), Problems in criterion-referenced measurements. U. of California.

Block, H. (1971). riterion-referenced measurements. School Review, 69, 289-298.

Carver, R.P. (1974). Two dimensions of tests: psychometric and edumeric. American Psychologists, 29, 512-518.

Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM Algorithm. Journal of Royal Statistical Society, 39, 1-39.

Ebel, R.L. (1971). Criterion-Referenced Measurements. School Review, 69. 282-288.

Hambleton, R.K. and Swaminathan, H. (1978). Criterion referenced testing and measurement: A review of technical issues and developments. Review of Educational Research, 48(1), 1-47.

Holland, P.W. and Wightman, L.E. (1982). Section pre-equating; A preliminary investigation. In P.W. Holland & D.B. Rubin (Eds), Test Equating. Academic Press.

Milman, J. (1973). Passing scores and test lengths or domain-referenced measures. Review of Educational Research, 43, 205-16.

Nitko, A.J.(1974). Problems in the development of criterion-referenced tests. In C.W. Harris, M.C. Alkin & W.J. Popham (Eds), Problems in criterion referenced measurement. University of California

Rubin, D.B. (1976). Comparing regressions when some predictor salues are missing. Technometrics, 1(2). 1-205.

Rubin, D.B. (1978). Relatiog tests sayer sychnctrika, 43(1), 1-39.

Thyer, D.T. (1982). Maximum likelihood estimation of the joint covariance matrix for sections of test given to distint samples with application to test equating. Educational Reseach Report 82-4. Eduational Testiong Service.

부록

번호	X_1	X_2	X_3	Y_2	Y_2	Y_1^*	Y_2^*
1	1	9.48	33.43	9.48	52.22	17.6988	40.7122
2	1	19.78	78.06	19.78	46.73	28.2825	32.8043
3	1	21.72	36.05	21.72	51.60	30.2760	39.8192
4	1	4.55	24.27	4.55	50.80	12.6330	38.6668
5	1	40.30	31.62	40.30	57.52	49.3677	48.3465
6	1	12.02	45.11	12.02	48.82	20.3088	35.8148
7	1	15.55	43.12	15.55	53.95	23.9360	43.2042
8	1	41.20	39.93	41.20	62.75	50.2925	55.8799
9	1	41.92	33.47	41.92	74.40	51.0323	72.6609
10	1	23.20	50.12	23.20	46.04	31.7967	31.8104
11	1	39.30	27.95	39.30	73.30	48.3402	71.0764
12	1	35.59	40.14	35.59	65.36	44.5280	59.6395
13	1	33.02	42.04	33.02	54.26	41.8872	43.6507
14	1	52.22	49.49	36.82	78.06	45.7919	77.9329
15	1	36.82	54.61	32.81	45.11	41.6714	30.4708
16	1	32.81	53.50	35.71	43.12	44.6513	27.6043
17	1	46.73	42.60	40.77	50.12	49.8507	37.6873
18	1	51.60	30.92	32.26	42.04	41.1063	26.0487
19	1	35.71	41.69	30.48	49.49	39.2772	36.7799
20	1	50.80	79.59	18.09	54.61	26.5460	44.1549
21	1	57.52	34.54	33.43	53.50	42.30852	42.5560
22	1	48.82	56.76	46.05	42.60	55.2761	26.8553
23	1	53.95	61.70	24.27	79.59	32.8962	80.1367
24	1	62.75	34.32	31.62	56.76	40.4486	47.2518
25	1	40.77	59.39	39.93	61.70	48.9875	54.3675
26	1	32.26	63.88	33.47	59.39	42.3496	51.0401
27	1	74.40	53.99	27.95	63.88	36.6775	57.5076
28	1	46.04	62.42	40.14	53.99	49.2033	43.2618
29	1	30.48	28.90	30.92	62.42	39.7294	55.4046
30	1	18.09	75.50	41.69	75.50	50.7960	74.2454
31	1	73.30	63.02	34.54	63.02	43.4491	56.2689
32	1	65.36	66.31	34.32	66.31	43.2230	61.0079
33	1	54.26	64.04	28.90	64.04	37.6537	57.7381