韓國數學敎育學會誌 < 數學敎育 >
1990. 12. 제 29 권 제 2 호 79 - 93

Journal of the Korea Society of Mathematical Education
Dec. 1990: Vol. 29, No. 2, 79 - 93

新·舊 두 考査 評價値 變換에 의한 眞分布와 母數 推定에 관한 硏究*

洪 錫强 (東國大學校)

1. 硏究의 必要性 및 目的

成功的인 學習 評價를 遂行하기 위하여는 質높은 學力 檢査 資料를 選擇 또는 製作 活用해야 한다. 좋은 檢査 資料는 두 가지 水準에서 고려될 수 있다.

그 하나는 어떤 試驗 問題 또는 學力 檢査 전체를 單位로 하는 경우이며, 다른 하나는 試驗 問題 또는 問項 하나 하나를 單位로 하는 경우이다. 前者의 경우 妥當度, 信賴度, 客觀度 등이 중요시되며 後者의 경우는 問項 困難度, 問項 辨別度, 問項 分布 등이 중요시 된다. 妥當度 (Validity) 에는 이를 判斷하는 準據 (Criterion) 의 種類에 따라

(가) 內容 妥當度 (Content Validity)

(나) 豫言 妥當度 (Predictive Validity)

(다) 公認 妥當度 (Concurrent Validity)

(라) 構因 妥當度 (Construct Validity)

가 있다. 信賴度 (Reliability) 에는

(가) 再檢査 信賴度 (Retest Reliability)

(나) 同形 檢査 信賴度 (Equvalent Form Reliability)

(다) 半分 檢査 信賴度 (Split-Half Reliability)

(라) 問項 內的 合致度 (Inter Item Consistency)

가 있으며, 다음 問項 分析을 위한 測度인 問項 良好度에는

(가) 問項 困難度 (Item Difficulty)

(나) 問項 辨別度 (Item Discrimination)

가 고려되어야 한다.

이 論文에서는 大學 入試 學力 考査가 高校 敎科 敎育 正常化에 미치는 영향이 큼을 고려하여 現行 高校 受驗生에게 施行되고 있는 模擬 考査, 기타 學力 考査가 위에서 고려한 모든 定義에 입각하여 良好한 評價 資料 인가를 檢討하고, 그 問項 反應 分布에서 眞分布 (True Score Distribution) 및 그 母數를 推定한 후 變換 解析 理論을 이용하여 新 舊 두 考査 評價値의 方程式을 算出하여 受驗生의 眞能力을 效率的이고 公正하게 評價할 수 있는 方法을 考案하고저 한다.

일반적으로 여러 學力 評價 硏究所에서 施行하고 있는 考査의 觀測値 (Observed Score) 의 總點이

* 本 硏究는 1989年度 韓國學術振興財團 硏究費 支援에 의한 結果임.

곧 受驗生의 能力으로 評價되고 特定한 出題 傾向에만 그들의 注意가 集中되어, 그런 類型의 問題 解決力 向上 [1, 2] 이 教科 教育의 目標에 到達된 것으로 認定되고 있는 現實이지만 본 研究 結果를 이용하여 受驗生의 眞能力의 分布를 구하고 위에서 고려한 모든 定義를 充足하는 良好한 問題가 더욱 많이 出題되므로서 高校 教科 教育의 正常化에 이바지 할 수 있으리라 생각된다.

다음은 同一한 母集團에서 抽出된 相異한 時期와 여러 回의 考査를 여러 個의 標本에서 試行할 때 考査의 難易度에 따른 評價値의 方程式化에 관한 研究는 Braun, H. I. 와 Holland, P. W. [4], Lord, F. M. [8] 등이 미국 Educational Testing Service 의 TOEFL, SAT 등의 考査에서 同形 檢査 信賴度 및 方程式化에 관한 研究를 시작으로

(1) 두 母集團의 合成

(2) 最良線型近似法의 $X = G(Y)\,F^{-1}(G(Y))$ 에 의한 變換評價値 算定

(3) (1) 에 의한 評價値로써 回歸 方程式의 線型性 檢定

등을 論하였고 특히 Lord, F. M. [5] 은 觀測値와 眞評價値의 相關性 檢定을 위한 數學的 模形의 研究에서

(1) 問項數가 增加, 減少할 때 觀測値 分布의 推定

(2) 百分率 等級化에 의한 두 考査 評價値에서 眞評價値의 方程式化

(3) 두 同形 檢査評價値에서 眞分布의 推定

(4) 不完全 Beta 分布型의 眞分布 推定 및 眞評價値의 方程式

등을 論하였다.

本 論文에서는 現行 各 學力 評價 研究所에서 이용하고 있는 標準化 點數 (T-Score) 와는 다른 새 評價法으로서 위에서 論한 同一한 또는 相異한 母集團에서 抽出한 相異한 數回의 時期와 相異한 數個의 標本에서 施行하는 評價의 假定下에서 二回 以上의 同型 檢査를 試行하고 妥當性의 定義 중 특히 出題 問項의 妥當度를 中心으로 同型 檢査 信賴度 및 각 問項의 良好度를 考察하고 負의 超幾何分布 (Negative Hypergeometric Distribution) 를 利用하여 觀測値의 眞分布型 및 母數를 推定하는 方法과 新 舊 두 考査 評價値의 方程式化에 의하여 受驗生의 眞能力을 公正하게 評價하는 方法을 提示하고저 한다.

2. 研究의 內容

2.1. 좋은 檢査의 選定

일반적으로 좋은 檢査의 경우에 內容 妥當度 및 信賴度를 檢定해야 하는 데, 이 때 信賴度 係數는 $\rho = 0.8$ 이상이면 바람직하며 동시에 좋은 問項의 選擇은 첫째, 檢査 實施의 必要性이나 目的을 分明히 해야 한다. 즉 受驗生들의 全般的인 學業 成就 水準을 알아보기 위한 檢査나 入學 試驗 또는 思考力 判斷, 能力의 辦別에 關心이 있는 경우 등에서 각각 問項의 選擇 基準이 달라야 한다. 둘째, 檢査 實施의 目標가 分明히 設定되었다면 이에 알맞은 教育 目標 二元 分類表가 준비되어야 한다. 셋째, 教育 目標 二元 分類表에 合當한 問項들이 選定되었다면 다음 問項 困難度 問項 辨別度 및 問項 反應 分布등 問項의 良好度 를 檢討해야 한다. 특히, 入學 試驗등과 같이 學生들의 能力을

엄격히 辨別해야 할 必要가 있을 경우에는 반드시 問項 困難 度係數를 고려해야 하며 기대하는 點數 分布에 따라 이에 알맞은 分布가 이루어 지도록 問項을 選擇해야 한다.

定義 1. 問項 g_i, $i=1, 2, \ldots, n$. 受驗生 a_i, $i=1, 2, \ldots, N$.
考査 점수를 X_{ai}, 妥當度 測定을 위한 準據(Criterion)
V_i, $i=1, 2, \ldots, n$라 할 때 妥當度(Validity) 係數는

$$\rho_{xv} = \frac{\sum_g \sigma_g \rho_{gv}}{\sum \sigma_g \rho_{gx}} \quad \text{이다. 단} \quad \rho_{gv} = \frac{\sum (g-\bar{g})(v-\bar{v})}{\sigma_g \sigma_v} \quad \text{임}$$

定義 2. Kuder-Richardson 21 의 信賴度 係數
問項 g_i, $i=1, 2, \ldots, n$가 二項 分布에 따르고 그 正答率을

$$P = \frac{\sum_{i=1}^{n} Pi}{n} \text{ 라 할 때 } \quad \alpha_{21} = \frac{n}{n-1}\left(1 - \frac{\bar{x} - \frac{\bar{x}^2}{n}}{\sigma_{x^2}}\right) \text{ 로 정의한다.}$$

定義 3. Horst P. 의 問項困難度指數

$$P_t = \frac{R_t - \frac{W_t}{K_i - 1}}{N_t - NR_t} \times 100$$

단 P_t : 問項困難度指數, R_t : 正答率
N_t : 受驗生 總數, NR_t : 各 問項의 未達項의 數
K_t : 答의 數 W_t : 誤答率

2.2. 非二項分布型 觀測値의 二項分布型 變換法

定義 4. 問項 g 의 觀測値 Y_g가 最大値 $Y_g=\alpha$와 最小値 $Y_g=\beta$를 가질때 二項變量

$$U_g = \frac{Y_g - \alpha}{\beta - \alpha} = 0 \text{ 또는 } 1 \text{ 이다.}$$

定義 5. 準據(Criterion) V가 標準正規分布 $N(0, 1)$에 따를 때

$$r_{gv} = \frac{M^+ - M^-}{S_v}\sqrt{p_g q_g}$$

를 Sample Point Biserial Correlation Coefficient 이라 한다.

단 $M^+ = \frac{1}{NP_g} \sum_{a=1}^{N} \sqrt[+]{a}$는 問項을 正答으로 表記한 受驗生의 $\sqrt[+]{a}$의 平均值.

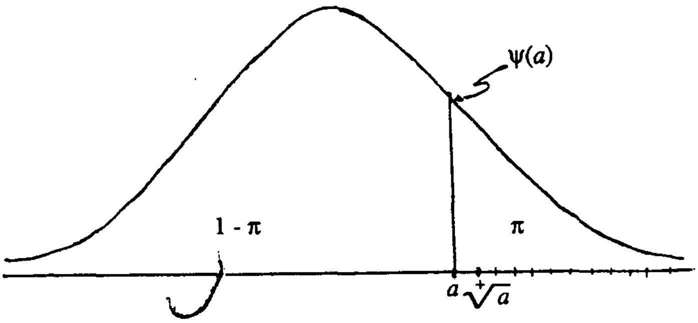
定義 6. 觀測値 X가 正規分布 $N(\mu, \sigma^2)$에 따르고 X上에서 準據 V의 回歸直線이 線型이고 $X<V$이면 $U=0$, $X>V$이면 $U=1$이라 하면 二項變量 U는

$$\pi = \Phi(V) = \int_V^{\infty} \Phi(X)\,dx, \quad \text{단} \quad \Phi(X) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$E(X \mid X \geq V) = \frac{\psi(V)}{\pi}, \qquad E(X \mid X < V) = -\frac{\psi(V)}{1-\pi} \text{ 이다.}$$

단, π는 각 問項의 困難度이며 $\varphi(V)$는 標準正規分布의 높이 (Ordinate) 이다.

[圖 1]



따라서 定義 4와 定義 6에서 妥當度 係數 (Biserial Correlation) 는

$$\rho'_{xv} = \frac{M^{+} - M^{-}}{\sigma_v} \quad \frac{\pi(1-\pi)}{\psi(V)}$$ 로 變換하여 計算한다.

Lord [8] 는 π의 函數로써 $\dfrac{\psi(V)}{\sqrt{\pi(1-\pi)}}$ 의 값을 다음과 같이 計算하였다.

定理 1. Biserial Correlation 과 Point Biserial Correlation 의 關係

$$\rho_{gv} = \frac{\psi(V_g)\ \rho'_{gv}}{\sqrt{\pi_g(1-\pi_g)}}$$

[表 1]

π	0.50	0.04 or 0.80	0.30 or 0.70	0.20 or 0.80	0.10 or	0.05 or
$\dfrac{\psi(V)}{\sqrt{\pi(1-\pi)}}$	0.798	0.79	0.76	0.70	0.58	0.47

2.3. 負의 超幾何分布型의 眞分布와 母數의 推定

二項誤差型 (Binomial Error Model) 에서 潛在力 (Latent Trait) 의 方程式

$$\delta(x) = \binom{n}{x} \int_0^1 g(\xi)\xi^{x}(1-\xi)^{n-x}$$

단, $\delta(x)$는 觀測値의 分布

$g(\xi)$는 眞分布의 未知分布에서

條件分布

$$P(\xi \mid x) = \frac{1}{\delta(x)}\ g(\xi)\binom{n}{x}\ \xi^{x}(1-\xi)^{n-x} \qquad (1)$$

이고 觀測値 x上에서 ξ의 回歸直線은

$$\mu_{\xi \mid x} = \frac{1}{\delta(x)}\binom{n}{x}\int_0^1 g(\xi)\xi^{x+1}(1-\xi)^{n-x}\ d\xi \qquad (2)$$

이다. 또 眞 評價値의 平均値는

$$\mu_{\xi|x} = 1 - \frac{n-x+1}{x} \frac{\delta(x-1)}{\delta(x)} \mu_{\xi|x-1}, \qquad x = 1, 2, \ldots, n \quad (3)$$

로 주어진다. 단, $\mu_{\xi|x} = \beta_{\xi x} x + \mu_{\xi}(1 - n\beta_{\xi x})$

$$\beta_{\xi x} = \frac{\sigma_{\xi} \rho_{\xi x}}{\sigma_x}, \quad \mu_{\xi|x-1} = \beta_{\xi x}(x-1) + \mu_{\xi}(1 - n\beta_{\xi x})$$

이때 眞 評價値의 回歸直線은

$$\delta(x)[1 - \beta_{\xi x} x - \mu_{\xi}(1 - n\beta_{\xi x})] = \frac{n-x+1}{x} \delta(x-1)[\beta_{\xi x}(x-1) + \mu_{\xi}(1 - n\beta_{\xi x})] \quad (4)$$

이다.

定理 2. 二項誤差型 (The Binomial Error Model) 에서 眞評價値가 觀測値 x 上에서 線型回歸直線을 갖는다면 觀測値의 確率分布는 負의 超幾何分布 [6] (Negative Hyoergeometric Distribution) 를 이룬다. 즉

$$h(x) = \frac{b^{[n]}}{[a+b]^{[n]}} \cdot \frac{(-n)_x (a)_x}{(-b)_x \, x!} \quad (5)$$

단, $n^{[x]} = n(n-1) \ldots (n-x+1)$

$(a)_x = a(a+1) \ldots (a+x-1)$

$$a = \left(-1 + \frac{1}{\alpha_{21}}\right)\mu_x, \qquad b = -a - 1 + \frac{n}{\alpha_{21}}$$

(證明)

(4)式에서 $\delta(x)$ 대신 $h(x)$를 代入하고 a와 b로써

$$h(x) = \frac{n-x+1}{x} \frac{a-1+x}{b+1-x} h(x-1)$$

의 漸化式에서 $x = 0, 1, 2, \ldots, n$일 때 定理는 自明하다.

定理 3. 負의 超幾何分布의 平均値와 分散은

$$\mu_x = \frac{na}{D}, \qquad \sigma_x^2 = \frac{na}{D} \frac{D-a}{D} \frac{D+n}{D+1} \quad (6)$$

이다. 단, $D = a + b - n + 1$.

定理 4. 眞評價値의 平均値와 分散은

$$\mu_{\xi} = \frac{\mu_x}{n}, \qquad \sigma_{\xi}^2 = \frac{1}{n^{[2]}}\left[\sigma_{x^2} - \frac{1}{n} \mu_x(n - \mu_x)\right] \quad (7)$$

이다.

2.4. 두 考査 評價値의 標本抽出과 方程式化를 위한 變換法

Levine R. S [7]. 은 同型檢査가 構造, 時間, 問項型 및 問項內容에서 同一性을 가지는 것이 바람직하

나 考査의 相對的 難易度에 따라 公正한 評價가 어려운 現實임을 지적하고 完全히 推定하기 어려운 여러 出題要因에 따라 最終 問項의 特性과 考査型의 變動이 必然的으로 수반되므로 新 舊 두 考査評價値의 方程式화(Equating)가 必要함을 강조하였다. Angoff W. H [3].는 相異한 型의 考査評價値의 方程式화를 위한 標本抽出을 다음과 같이 假定하였다.

[表 2]

	X	Y	V
標本 $P \to P_1$	觀測可能	觀測不能	觀測可能
標本 $Q \to Q_1$	觀測不能	觀測可能	觀測不能

	X	Y
標本 $P \to P_1$	觀測可能	觀測不能
標本 $P \to P_2$	觀測不能	觀測可能

두 考査 眞 評價値의 x, y의 分布函數를 $F(X)$, $G(Y)$라 할때 母集團 P의 標本에서 $F(X) \le G(Y)$이면 難易面에서 X는 Y보다 쉬운考査이며 그 標本抽出은 表 2]의 경우로 假定한다.

定義 7. $X^* \le x$에 대한 P에서 應試者의 比率을 $F^*(x) = G(e_x^{-1}(x))$라 할 때 $F \neq F^*$인 母集團 P가 存在하면 X, Y두 考査는 P上에서 e_x에 의한 方程式化가 不可能하다. 實際로 觀測値의 分布가 離散型으로 計算되므로 $F = F^*$인 경우는 不可能하지만 分布函數의 메트릭(Metric)에서 $F \le F^*$ 또는 $F^* \le F$이면 方程式化가 可能한것으로 假定하여 다음과 같이 定義할 수 있다.

定義 8. X, Y 두 考査는 $F = F^*$이면 e_x에 의하여 P에서 變換된다.

定理 5. 만일 두 分布가 同等百分率에 의하여 $F = F^*$이면 $e_x(Y) = F^{-1}(G(Y))$

즉 $e_x(Y) = aY + b$, $F^{-1}(G(Y)) = aY + b$

$$F(x) = H\left[\frac{x - \mu_x}{\sigma_x}\right], \qquad G(Y) = \left[\frac{y - \mu_y}{\sigma_y}\right]$$

단, μ_x, μ_y, σ_x^2, σ_y^2는 X, Y 두 考査의 眞平均値와 分散임

$$a = \frac{\sigma_x}{\sigma_y}, \qquad b = \mu_y - \frac{\sigma_x}{\sigma_y}\mu_y$$

$$X^* = e_x(Y) = \mu_x + \frac{\sigma_x}{\sigma_y}(Y - \mu_y) \qquad (8)$$

이다.

(8)式은 두 考査가 相互從屬인 경우 ρ_{xy}가 고려된 일반 回歸直線과 달리 同型檢査이므로 $\rho_{xy} = 1$로 나타낸다.

3. 計算 例 및 檢討

3.1. 標本抽出과 妥當性 計算

[表 3]과 같이 考査 平均値 標本을 抽出하기위해 서울市 所在 각 入試學院 學力評價硏究所의 考査資料를 利用하였다.

[表 3]

	學力評價研究所		
	J	D	C
$P \rightarrow P_1$	K 校	I 校	K 校
$P \rightarrow Q_1$	J 校	觀測不能	J 校

	學力評價研究所	
	D	C
$P \rightarrow P_1$	K 校	S 校
$P \rightarrow P_2$	U 校	C 校

다음의 例는 많은 科目 중 學生數 $N = 56$, 問項數 $n = 33$ 인 數學考査의 觀測値써 資料處理 한 結果이며 社會科目의 N가 增加 한 경우의 圖表는 [圖 5] 에 수록하였다.

[表 4] 受驗生의 成績 資料

채 점 표 (문항별)					
번 호 / 과 목	1 2 3 4 5 6 7 8 9 0 (1)	1 2 3 4 5 6 7 8 9 0 (2)	1 2 3 4 5 6 7 8 9 0 (3)	1 2 3 4 5 6 7 8 9 0 (4)	1 2 3 4
국어 I, 한문 I, 국사	1× 1 ×× 1 1 1 1 1	1 1 1 1 1 1 1 1××	1 1 1×1 1 1 1 1 1	×1 ×1 1 1 ×1×	3 2 2 2
수 학	2 2 2 2 2 2 2 2 2 2	2 2 2 × × 2 2× 2 2	× 2 2 2 2 2	3 ×3 3 3 4 4	완전한 표는 P.93 [표4] 를 볼것.
사 회	1 1 1 1 1 1 1 1 1×	×1 1 1	1 2 1		
영어, 실업, 제2외국어	1 × × ×1 1 1 1 1 1	×1 ×1×1 1 1 × ×1	1 1 1 1 1 ×1 1 1	1 ×1× ×1 1 1 1×	
국민윤리, 국어 II	1 ×1 1 1 ×1 1 1 1	×1 ×1×1 1×1 2 2	2		
과학, 예, 체능	1 ×1 1 1 1 1 1 1 1	1 1 1 1 1 2 1 1	×1 1×1 1 1 1 1 1 1	1× × 2 1	

(1) 主觀式 評價値의 二項變量化 變換

$$U_{41} = \frac{Y_g - \alpha}{\beta - \alpha} = \frac{3 - 0}{3 - 0} = 1$$

(2) 問項困難度 計算

$$Pt_{41} = \frac{R_t - \dfrac{W_t}{K_i - 1}}{N_t - NR_t} \times 100 = 0.64285$$

(3) Kuder-Richardson 21 의 信賴度係數 計算

$$\alpha_{21} = \frac{n}{n-1}\left[1 - \frac{\bar{x} - \dfrac{\bar{x}^2}{n}}{Sx^2}\right] = \frac{33}{33-1}\left[1 - \frac{17.160720 - \dfrac{(17.160720)^2}{33}}{36.92647}\right] = 0.76276$$

(4) $P_{41} = 0.64285$ $Sx_{41}{}^2 = p_{41}q_{41} = 0.237249$

$$M^+ = E[v/u_g = 1] = \frac{-0.37 - 0.2737 - 0.1774 - \cdots\cdots + 2.9037 + 3.0000}{36} = 1.315000$$

$$S^2_{M^+} = 1.00480$$

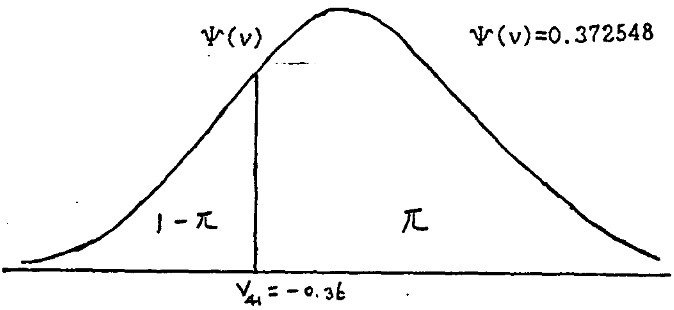
$$M^- = E[v/u_g = 0] = \frac{-0.30000 - 2.8685 - \cdots\cdots - 0.5015}{20} = -1.750750$$

$$S^2_{M^-} = 0.574967$$

$$M = \frac{-0.37 \cdots\cdots + 3.0000}{56} = 0.220089$$

$$S_{M^2} = 3.006410$$

[圖 2]



標準正規分布 數値로써 N의 數가 크면 連續變量으로, 작으면 離散變量으로 增分을 정해 實數値를 計算 함.

(5) 妥當度係數 計算檢算

$$\rho'_{xv} = \frac{M^+ - M^-}{\sigma_v}\left[\frac{\pi(1-\pi)}{\psi(v)}\right] = 0.99999$$

$$\rho_{gv} = \frac{\psi(v_g)\rho'_{gv}}{\sqrt{\pi_g(1-\pi_g)}} = 0.8470276$$

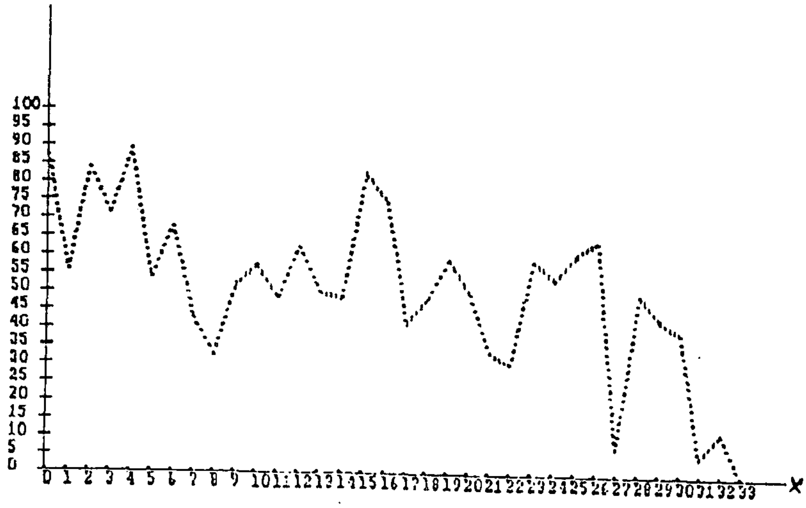
즉 問項 41은 妥當性이 높은편이며 各 問項別 標本妥當度係數γ_{gv}와 Point Biserial Correlation Coefficient ρ_{gv}은 다음 表 5]와 같다. 이結果에서 두 係數는 N가 작은 경우이지만 거의 유효숫자 소숫점 3째 자리까지 일치하고 있고 標本의 크기 N와 n가 클 때는 母妥當度係數値를 더욱 정확히 計算할 수 있다.

[表 5]

問項		γ_{gv}	ρ_{gv}
1	1	0.666418	0.666417
	2	0.863428	0.863335
	3	0.720231	0.720592
	4	0.820700	0.820871
	5	0.632313	0.631792
	6	0.864800	0.867346
	7	0.835725	0.835381
	8	0.861809	0.861935
	9	0.835987	0.835712
	10	0.865712	0.865699
2	1	0.861579	0.861680
	2	0.865739	0.865733
	3	0.851850	0.851851
	4	0.865987	0.865986
	5	0.865029	0.865023
	6	0.739898	0.740545
	7	0.800450	0.800449

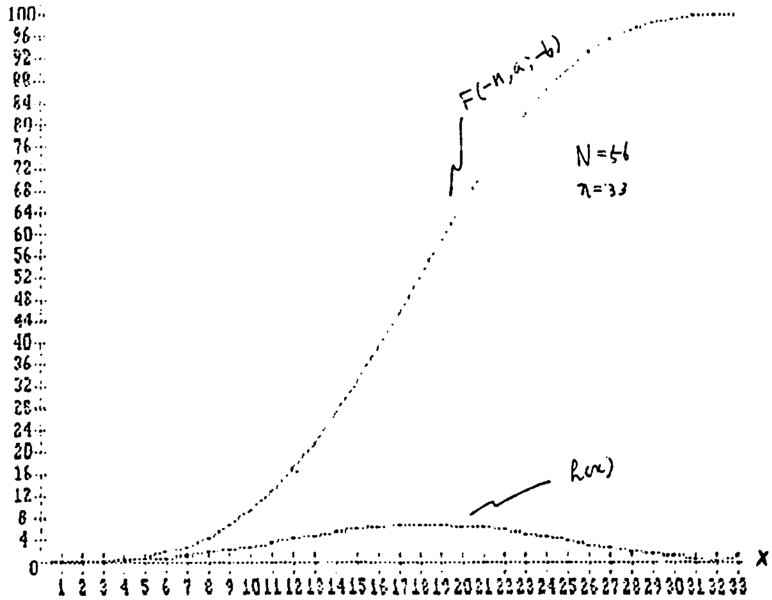
	8	0.859358	0.859434
	9	0.865731	0.865726
	10	0.859129	0.859192
3	1	0.865731	0.865985
	2	0.842308	0.842188
	3	0.829491	0.829876
	4	0.859129	0.859123
	5	0.864800	0.864762
	6	0.855835	0.855874
4	1	0.847134	0.847628
	2	0.528949	0.527528
	3	0.865987	0.865986
	4	0.861809	0.861936
	5	0.856130	0.856207
	6	0.467782	0.469840
	7	0.661264	0.661263

[圖 3]



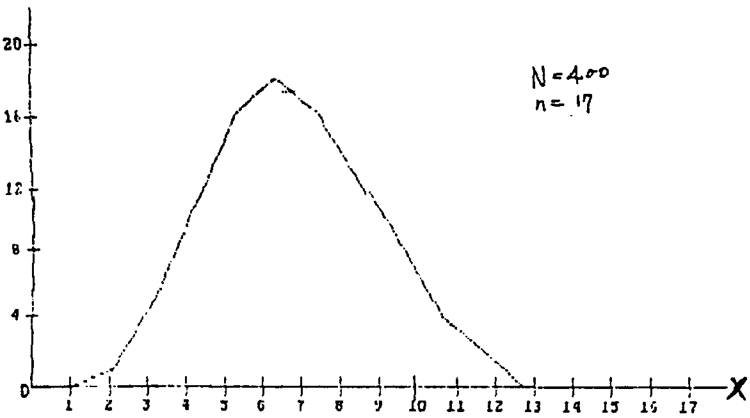
數學科目 의 正答率 分布

[圖 4]



[表 6]의 確率分布와 累積分布 그래프

[圖 5]



[表 8]의 確率 分布

3.2. 負의 超幾何分布型 眞分布近似值. 母數計算 및 最終評價

[表 6] 眞分布의 確率函數值와 累積分布函數值

(단위 %)

問項	x	$h(x)=\dfrac{b^{(n)}\ (-\Sigma)_x(a)_x}{[a+b]^{-(n)}\ (-b)_x\ x!}$	$F(-n,a:-b)=\sum_{x=0}^{33}h(x)$
1	0	0.008556	0.008556
	1	0.040811	0.049367
	2	0.115187	0.164554
	3	0.250055	0.414609
	4	0.460885	0.875495
	5	0.758062	1.633557
	6	1.145447	2.779004
	7	1.619671	4.398675
	8	2.170105	6.568781
	9	2.779435	9.348215
	10	3.424714	12.772930
2	1	4.078813	16.851740
	2	4.712124	21.563870
	3	5.294411	26.858280
	4	5.796695	32.654970
	5	6.193060	38.848030
	6	6.462284	45.310320
	7	6.589205	51.899520
	8	6.565747	58.465270
	9	6.391547	64.856820
	10	6.074138	70.930950
3	1	5.628667	76.559620
	2	5.077134	81.636760
	3	4.447180	86.083940
	4	3.770453	89.854390
	5	3.080625	92.935010
	6	2.411142	95.346160
4	1	1.792838	97.138990
	2	1.251542	98.390540
	3	0.805880	99.196420
	4	0.465476	99.661890
	5	0.229816	99.891710
	6	0.88068	99.979780
	7	0.020226	100.000000

定理 4. 에 의하여 $\bar{X} = 17.16072$, $\alpha_{21} = 0.76276$, $a = 5.337475$, $b = 36.92647$, $D = 10.26394$ 로써 $\mu_x = 17.160723$, $\sigma^2_x = 34.71913$, 또 $\mu_\xi = 0.5200219$, $\sigma^2_\xi = 0.0250779$ 이므로 正答率 $\mu_\xi = 0.5$ 와 $\sigma^2_\xi = 0.025$ 와 實驗式이 一致함을 알 수 있다.

다음 表 7]은 각 受驗生의 點數와 眞評價值를 比較한 數值이다.

[表 7]

番號	觀測値	眞評價値
1	85.33333	98.39054
2	38.66666	26.85828
3	37.33333	32.65497
4	41.33333	32.65497
5	82.66666	98.39054
6	60.00000	76.55962
7	32.00000	21.56387
8	26.66666	9.34821
9	53.33333	64.85682
10	54.66666	70.93095
11	92.00000	99.66189
12	24.00000	9.34821
13	68.00000	86.83940
14	66.66666	81.63676
15	66.66666	64.85682
16	66.66666	81.63676
17	22.66666	6.56878
18	33.33333	21.56387
19	45.33333	51.89952
20	37.33333	32.65497
21	37.33333	32.65497
22	53.33333	58.46527
23	54.66666	64.85682
24	44.00000	38.84803
25	38.00000	32.65497
26	49.33333	51.89952
27	62.66666	81.63676
28	21.33333	6.56878
29	52.00000	58.46527
30	64.00000	81.63670
31	58.66666	76.55962
32	60.66666	76.55962
33	45.33333	45.31032
34	29.33333	16.85174
35	38.66666	32.65497
36	57.33333	76.55962
37	93.33333	99.89171
38	85.33333	98.39054
39	62.66666	12.77293
40	64.00000	81.63676
41	25.33333	9.34812
42	52.00000	51.89952
43	30.66666	16.85174
44	49.33333	51.89952
45	48.00000	51.89952
46	52.00000	58.46527
47	26.66666	12.77293
48	56.00000	70.93095
49	38.66666	26.85828
50	50.66666	58.46527
51	29.33333	16.85174
52	40.00000	32.65497
53	69.00000	86.08394
54	34.66666	26.85828
55	76.00000	95.34616
56	21.33333	6.56878

[表 7]에서 例로 14와 15번 觀測値가 같은 66.6666이지만 難易度가 반영된 眞評價値는 각각 다른데 이것은 어려운 問題를 解決한 學生이 더욱 高得點을 한 結果를 보이고 있다. 다음 [表 8]은 두 社會科目의 眞分布 및 確率函數의 回歸方程式値이며 (8)式에 의하여 2회 考査確率函數數値를 代入하여 最終 成績인 眞評價値를 計算한 結果이다.

[表 8] 社會科目 의 眞分布, 回歸方程式値 및 眞 評價値 (단위 %)

數値名 / 問 項	h(x)	h(y)	回歸方程式値	眞評價値
0	0.472858	0.174332	0.381633	0.381633
1	2.787988	1.154457	0.699754	1.081387
2	7.969395	3.815564	3.635796	4.717183
3	14.662000	8.347365	8.635803	13.352986
4	19.437690	13.5211340	14.344330	27.697316
5	19.697690	17.190350	18.392410	46.089726
6	15.794730	17.744040	19.003300	65.093026
7	10.232780	15.175190	16.169050	81.262076
8	5.420900	10.877390	11.427220	92.689296
9	2.360954	6.567226	6.6717748	99.361044
10	0.845116	3.337523	3.108365	102.469409
11	0.247152	1.418826	0.991437	103.460846
12	0.058313	0.498114	- 0.024399	103.436447
13	0.010864	0.141328	- 0.418047	103.018400
14	0.001543	0.031284	- 0.539460	102.478940
15	0.000157	0.005089	- 0.568361	101.910579
16	0.000010	0.000543	- 0.573377	101.447202
17	0.000000	0.000029	- 0.573944	100.873258

[表8] 에서 眞評價値 가 100 % 를 초과한 結果 는 眞分布型 資料適合 에서 派生 된 傳播 誤差 로 인한것이며 正答數 에 따라 最終 眞評價値 를 計算 하고 受驗生 의 實觀測値 와 比較 할 수 있다.

4. 結論

이 論文 에서 抽出 한 學生數 나 問項數 의 크기가 작고 各 學力評價研究所 의 出題形式 에 따라 약간 相異 한 研究 結果 가 나올것이 豫想 되나 대체로 研究 結果 및 提言 을 要約 하면 다음과 같다.

1. 現行 學力評價研究所 의 出題 問項 의 妥當度 및 信賴度係數値 가 不安定 하다.
2. 高相關 信賴度 의 問項 을 出題 하기 위해서 教科目別 및 各 教科目內 에서 問項 採點 配分 을 더욱 細分化 시킬 필요가 있다. 例 로써 國語, 英語, 數學 의 必須 科目 과 選擇 科目 의 配点 이 다르므로 評價値 에서 重要 科目 의 影響 을 많이 받고 있고, 또 國語, 英語, 기타 暗記 科目 에서는 問項數 가 많지만 數學 에서는 問項數 가 작으며 뿐만아니라 各 科目 의 主觀式 問項 에서 配点 을 더욱 細分化 시킬 必要 가 있다. 즉, 問項數 가 많고 配点率 이 細分化 된 科目 에서는 觀測値 의 二項變換 의 效率性 을 더욱 높일 수 있다.
3. 各 教科目 에서 出題 資料 抽出 範圍 를 더욱 넓혀 既存 出題 問項 만 重要視 하는 폐단을 시정하고, 良好 한 評價 資料 및 多樣 한 客觀式 技法 을 導入 하여 全科目 에서 內容 의 均等 한 學習 을 試圖 해야 한다.

4. 問項 配點의 細分化와 評價 結果의 散布度를 크게하므로써 受驗生의 個人差를 엄격히 辨別하기위한 評價의 目的에 부합되게 各 類型別 思考力 深化 能力을 檢定할 수 있는 評價 資料를 出題해야 한다.

5. 이 研究 結果에서 例示한 評價法이 問項의 難易度 및 信賴度가 반영된 學生의 眞 能力을 評價하는데 現行 評價法 (T-Score) 보다 더 效率性이 높다.

6. 受驗生數 와 問項數 가 크고 妥當度 및 信賴度 係數가 크면 眞分布型의 安定性이 더욱 커지므로 기대하는 評價分布를 豫想하기 쉽다.

7. 新 舊 두 考査 觀測値를 眞評價値로 變換한 후 두 評價値의 方程式化의 結果로써 出題 時期와 出題 形式에 左右되지 않는 能力의 眞分布를 推定할 수 있다.

8. 이 研究에서 開發된 프로그램을 資料 處理에 利用할 수 있다.

參考 文獻

1. 洪 錫强 : 中等學校의 數學的 思考力 增進을 위한 效果的인 數學指導法의 試案, 韓國數學敎育學會誌 第26卷 第2號 (1988), 15-41.

2. 洪錫强 : 數學敎育에서 理解力 深度의 測定과 方法, 敎育問題研究 (東國大學校 敎育問題研究所) 第五輯 (1988), 83-95.

3. Angoff, W.H. : *Scales, Norms and Equivalent Scores*, in "Educational Measurement" (Ed. by Thorndike R. L.), Washington, American Council on Education, 1971, pp. 508-600.

4. Braun, H. I. and Holland, P.W. : *Observed Score Test Equating*, in "A Mathematical Analysis of Some ETS Equating Procedure, Test Equating" (Ed. by Holland P.W. and Rubin D. B.), Educational Testing Service. 1982, pp. 9-55.

5. Keats, J.A. and Lord, F.M. : *A Theoretical Distribution for Mental Test Score*, Psychometrika **27** (1962) pp. 59-72.

6. Kendall and Stuart : "The Advanced Theory of Statistics", Vol. 1, Distribution Theory, New York, Hafner , 1958, pp. 134.

7. Levine, R.S. : "Equating the Score Scales of Alternate Forms Administered to Samples of Different Ability RB- 55-23", Educational Testing Service, 1955.

8. Lord, F.M. : "Applications of Item Response Theory to Practical Testing Problems", LEA, 1980, pp. 235-264.

ABSTRACT

In this thesis the following studies have been tried :

1. To estimate reliability and validity of the items of scholastic achievment tests that had been tested by the evaluation service centers.
2. To smooth the sample frequency distribution of observed scores and to estimate the frequency distribution of observed scores approximating to the Negative Hypergeometric Distribution.

3. To equate two tests and to estimate item-test regressions, item response functions and their parameters to evaluate correctly and efficiently the scholatic ability of individual student.

[表 4] 受驗生의 成績 資料

채 점 표 (문항별)								
번 호 / 과 목	1 1234567890	2 1234567890	3 1234567890	4 1234567890	5 1234567890	6 1234567890	7 1234567890	8 1234567890
국어 I, 한문 I, 국사	1×1××11111	11111111××	111×111111	×1×111×1×	32222×	×111111111	11111111×	21×
수 학	2222222222	222××22×22	×22222	3×333444				
사 회	111111111×	×111	121					
영어, 실업, 제 2외국어	1×××1111111	×1×1×111××	11111111×11	11×1××1111	×11	××223223	111111×11	1×111222
국민윤리, 국어 II	1×111×1111	×1×1×11×12	22					
과학, 예, 체능	1×111111111	1111 1211	×11×111111	11× ×21				