看護専門科目Ⅰ・Ⅱにおける異なるテストの得点等化

伊藤圭（大学入試センター），柳井晴夫（聖路加看護大学）

1 はじめに

2009年度に行われた看護系大学共用試験（CBT）のモニター調査では，3つの受験者グループに対して，基礎医学，看護専門科目Ⅰ，看護専門科目Ⅱのテストを表1に示すとおり割り当てた。看護専門科目Ⅰおよび看護専門科目Ⅱではそれぞれ3種類のテストが作成された。これは，一回の調査で出題することができる問題項目数に限度があるため，問題項目を複数のテストに振り分け，各テストを異なる受験者グループに割り当てる必要があったためである。しかしながら，本研究で開発する共用試験はCBTであるため，同一の特性を測りながらも出題項目が異なる複数のテストが実際に存在することになる。また，異なるテスト間の得点が比較可能であるかどうかもCBTを開発する上で本質的な問題である。

本節では，全てのグループが受験している基礎医学（テストA）の成績を通してグループ間の学力差を考慮した上で，看護専門科目Ⅰにおける複数のテストB1，B2，B3を等化して難易度比較を行う。また，同様にして看護専門科目Ⅱにおける複数のテストC1，C2，C3の難易度比較も行う。

表1　受験者グループとテストの割り当て

グループ	基礎医学	看護専門科目Ⅰ			看護専門科目Ⅱ		
	A	B1	B2	B3	C1	C2	C3
1	○	○			○		
2	○		○			○	
3	○			○			○

2 共通テストを用いた等化

表1から分かるように，基礎医学（テストA）は全てのグループが受験しているため，いわゆる共通テストを用いた等化法が利用できる。ここでは，等化対象テストの共通テストへの回帰を利用したTucker法による等化を行った。

今，受験者グループ f がテスト X を，g がテスト Y を受験し，さらに両グループが共通テスト Z を受験しているとする。一般にグループ k（人数 N_k）のテスト W の得点を $w_{k,i}$（$i=1,2,\ldots,N_k$），その平均点および標準偏差をそれぞれ $\bar{x}_k$，s_{x_k} と表記することにすると，各グループの各テストに対する得点，平均点，標準偏差は表2の通りとなる。但し，グループ h は f と g を合わせた集団である。ここで，実際の調査から直接得られるものは表2の太枠以外の部分のデータであり，これらのデータに基づき，太枠内，すなわちグループ h のテスト X および Y の平均点と標準偏差を推定することになる。4つの量 $\bar{x}_h$，s_{x_h}，$\bar{y}_h$，s_{y_h} が求められれば，グループ h をテスト X および Y の同時受験者集団とみなすことができるため，テスト X と Y の難易度を平均点 $\bar{x}_h$，$\bar{y}_h$ によって直接比較することができる。さらに，テスト Y の得点 y を線形変換

表 2 共通テストデザインにおける各グループの受験状況と統計量

グループ	テストX	テストY	共通テストZ
f	x_{fi}，$\bar{x}_f$，s_{x_f}	受験していない	z_{fi}，$\bar{z}_f$，s_{z_f}
g	受験していない	y_{gi}，$\bar{y}_g$，s_{y_g}	z_{gi}，$\bar{z}_g$，s_{z_g}
h（f & g）	$\bar{x}_h$，s_{x_h}	$\bar{y}_h$，s_{y_h}	z_{hi}，$\bar{z}_h$，s_{z_h}

$$y' = \frac{s_{x_h}}{s_{y_h}}\left(y - \bar{y}_h\right) + \bar{x}_h$$

によって新たな得点y'に置き換えれば，グループhにおけるテストXとYの平均点と標準偏差を一致させることができる（$\bar{x}_h = \bar{y}'_h$，$s_{x_h} = s_{y'_h}$）。上式は“テストYをテストXへ等化する”変換式で，テストYの得点をテストXの得点尺度上で表すことにより，両テストの得点を比較可能にするものである。

Tucker 法では，$\bar{x}_h$，s_{x_h}，$\bar{y}_h$，s_{y_h}を推定するために，共通テストZの得点zからテストXの得点xおよびYの得点yを予測する線形回帰分析を考える。今，xのzへの回帰のみに注目すると，グループfおよびhのそれぞれに対して，

$$x_{f,i} = \alpha_{x_f z_f} + \beta_{x_f z_f} z_{f,i} + \varepsilon_{x_f z_f, i} \quad , \quad x_{h,i} = \alpha_{x_h z_h} + \beta_{x_h z_h} z_{h,i} + \varepsilon_{x_h z_h, i}$$

となる。このとき，xとzの平均と標準偏差，および誤差分散の間には

$$\bar{x}_f = \alpha_{x_f z_f} + \beta_{x_f z_f} \bar{z}_f \quad , \quad \bar{x}_h = \alpha_{x_h z_h} + \beta_{x_h z_h} \bar{z}_h \,,$$

$$\left(s_{x_f}\right)^2 = \beta_{x_f z_f}{}^2 \left(s_{z_f}\right)^2 + \left(\sigma_{x_f z_f}\right)^2 \quad , \quad \left(s_{x_h}\right)^2 = \beta_{x_h z_h}{}^2 \left(s_{z_h}\right)^2 + \left(\sigma_{x_h z_h}\right)^2$$

という関係がある。ここで，$\left(\sigma_{x_f z_f}\right)^2$，$\left(\sigma_{x_h z_h}\right)^2$は誤差$\left(\varepsilon_{x_f z_f, i}\right)$，$\left(\varepsilon_{x_h z_h, i}\right)$の分散を表す。

Tucker 法では，さらに，切片α，回帰係数β，誤差分散σ^2がグループfとhで等しいという仮定を置く。すなわち，$\alpha_{x_f z_f} = \alpha_{x_h z_h}$，$\beta_{x_f z_f} = \beta_{x_h z_h}$，$\left(\sigma_{x_f z_f}\right)^2 = \left(\sigma_{x_h z_h}\right)^2$である。上で述べた$x$と$z$の平均と標準偏差，および誤差分散の関係式にこの仮定を適用することにより，

$$\bar{x}_h = \beta_{xz}\left(\bar{z}_h - \bar{z}_f\right) + \bar{x}_f \quad , \quad s_{x_h} = \sqrt{\beta_{xz}{}^2 \left\{ \left(s_{z_h}\right)^2 - \left(s_{z_f}\right)^2 \right\} + \left(s_{x_f}\right)^2}$$

が成り立つことが分かる。また，yのzへの回帰についても同様に考えることにより，

$$\bar{y}_h = \beta_{yz}\left(\bar{z}_h - \bar{z}_g\right) + \bar{y}_g \quad , \quad s_{y_h} = \sqrt{\beta_{yz}{}^2 \left\{ \left(s_{z_h}\right)^2 - \left(s_{z_g}\right)^2 \right\} + \left(s_{y_g}\right)^2}$$

が成立する。この結果からグループhに関する推定値$\bar{x}_h$，s_{x_h}，$\bar{y}_h$，s_{y_h}を求め，テストYをテストXへ等化する変換式を得ることができる。以上が Tucker 法による等化の概要である。なお，等化に関する詳細については，柳井，前川（1999），Kolen，Brennan（2010），Linn（1989）[池田，藤田，柳井，繁桝（1992）]を参照されたい。

3 結果と考察

前節で述べた等化方法を用いてテストB2，B3をテストB1へ，テストC2，C3をテストC1へ等化した。テストB1，B2，B3の共通テストAへの回帰における回帰係数βの値は，それぞれ0.6333，0.5951，0.6432，テストC1，C2，C3の共通テストAへの回帰における回帰係数βの値は，それぞれ0.7192，0.6962，0.6534であった。

表 3 は各受験者グループのテスト成績を表したものである。各グループにおいて，上段の数値が平均点，下段の数値が標準偏差である。太枠で囲まれている部分は，Tucker 法における仮定に基づいて推定された値であり，実際の調査から直接得られるものではない。太枠で囲まれた部分以外は実際の調査で得られた受験者の解答結果を集計したものである。また，表中の網掛け部分が等化結果であり，平均と標準偏差は，B2→B1 の等化で 103.20，14.35（グループ2におけるB1の結果），B3→B1の等化で102.42，13.78（グループ3におけるB1の結果）,C2→C1で89.04,13.01(グループ2におけるC1の結果),C3→C1で89.22，12.31（グループ 3 における C1 の結果）となった。共通テストである基礎医学の成績を見るとグループ1がグループ2，3に比べてやや学力が高く，2と3はほとんど同じ学力であることが分かる。看護専門科目 I および II における等化後の結果においても同様の傾向が見られ，共通テストである基礎医学のグループ間の学力差が適切に等化に反映した結果となっていることが分かる。さらに，看護専門科目 I および II におけるグループ別の相対累積度数分布を図 1 および図 2 に示した。等化前にはグループ間で得点分布に違いが見られるが，等化後は各グループが非常に似た分布を有していることが分かる。

看護専門科目 I における3つのテスト（B1，B2，B3）の難易度，および看護専門科目 II における3つのテスト（C1，C2，C3）の難易度を比較するには，グループ（1+2）およびグループ（1+3）の成績を比較すれば良い。B1の平均点はB2より6.02点高く，B3より2.22

表 3 平均点および標準偏差

グループ		基礎医学	看護専門科目 I			看護専門科目 II		
		A	B1	B2	B3	C1	C2	C3
1	Mean	74.56	104.65	―	―	90.54	―	―
	SD	11.58	14.04	―	―	12.64	―	―
2	Mean	72.40	103.20	97.31	―	89.04	88.92	―
	SD	12.38	14.35	12.15	―	13.01	13.38	―
3	Mean	72.70	103.42	―	101.24	89.22	―	91.61
	SD	10.83	13.78	―	12.87	12.31	―	11.55
1+2	Mean	73.59	104.03	98.01	―	89.84	89.74	―
	SD	12.00	14.18	12.01	―	12.84	13.21	―
1+3	Mean	73.67	104.09	―	101.87	89.90	―	92.25
	SD	11.27	13.94	―	13.02	12.50	―	11.72

点高い。C1 と C2 の平均点差は 0.1 でほとんど同じ難易度であり,C1 の平均点は C3 より 2.35 点低い。したがって，看護専門科目 I では B2 が比較的難しいテストであり，看護専門科目 II では C3 が比較的易しいテストであったことが分かる。

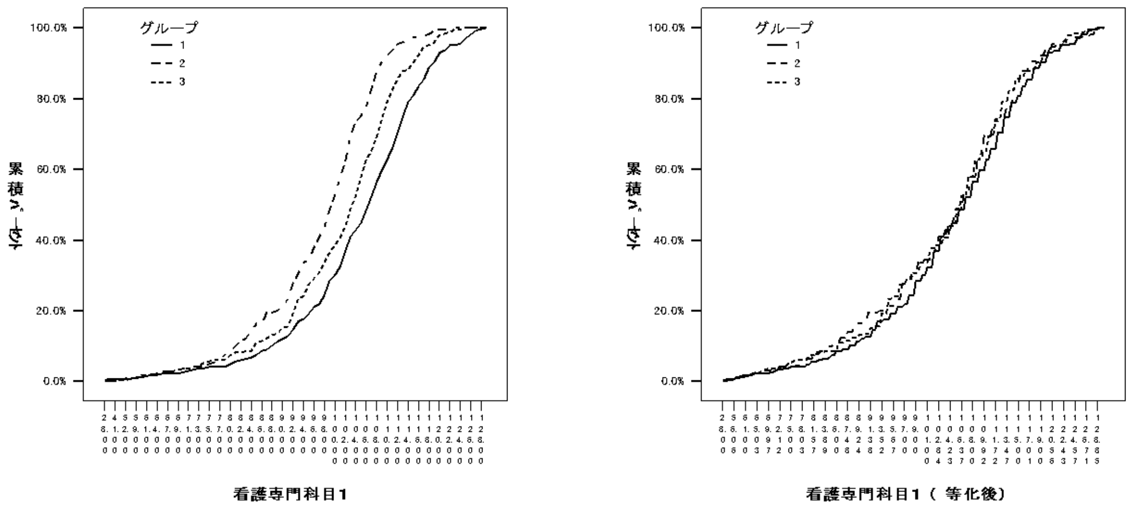

図 1 看護専門科目 I におけるグループ別の相対累積度数分布（左：等化前，右：等化後）

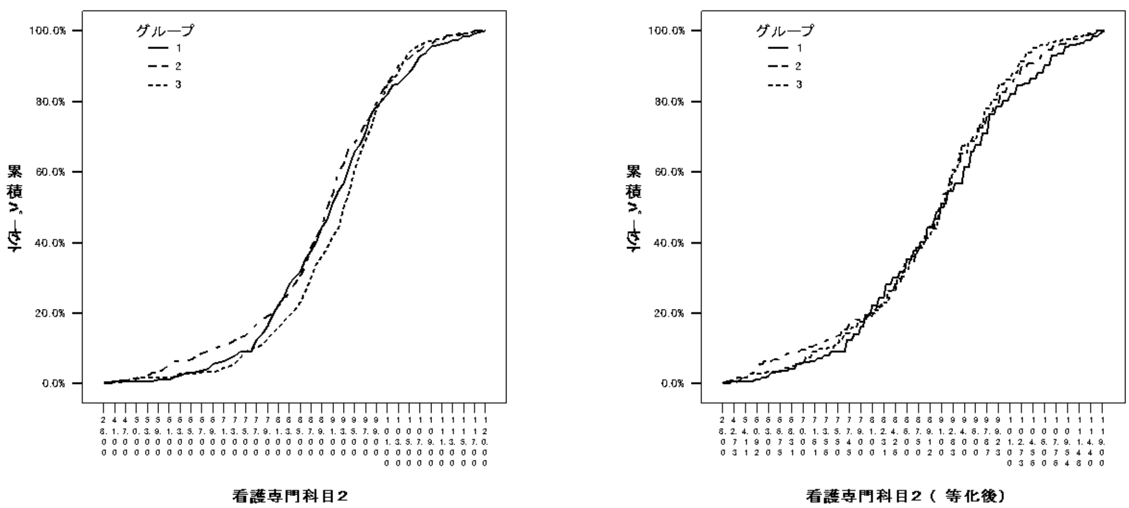

図 2 看護専門科目 II におけるグループ別の相対累積度数分布（左：等化前，右：等化後）

参考文献

柳井晴夫，前川眞一［編］(1999)．大学入試データの解析［理論と応用］．現代数学社.

Kolen, M. J. and Brennan, R. L. (2010). *Test Equating, Scaling, and Linking: Methods and Practices* (2nd ed.). Springer Science+Business Media.

Linn, R. L. (Ed.) (1989). *Educational Measurement* (3rd ed.). American Council on Education and Macmillan. ［池田央，藤田恵璽，柳井晴夫，繁桝算男（訳編）(1992). 教育測定学第 3 版（上下） C.S.L. 学習評価研究所.］