広島大学大学院工学研究科研究報告，60，1 (2011)

Bulletin of the Graduate School of Engineering, Hiroshima University. Vol.60, No.1, 2011

試験の得点分布モデルとその変換

樋口　勇夫*

The model of experimental score and transformation

Isao HIGUCHI

When explaining an educational effect using an achievement test, the examination from which difficulty differs to the group from whom an examinee differs will be done. In such a case, although it is thought that the difference in difficulty comes out in the average mark, the method for the difference reflects has not been established yet. In this paper, the model which transformations into another random variable on the random variable is used, and which takes the value of the limited section is based on the idea of drawing the percentage of correct answers in that examination using the difficulty of an examinee's potential academic ability and an examination. The method of estimating the parameter showing the difference in difficulty is raised using this model.

Key words: experimental score, achievement test, stochastic model, transformation, moment estimation

1. はじめに

一般に学力試験を行う場合、その試験での得点が重要であり、他の試験との難易度差が取りざたされることは多くはない。しかし、大学入試センター試験では科目間の平均点に20点以上の差が出た場合は得点調整が行われる可能性がある。また、工学系数学統一試験(EMaT)の結果をもとに教育効果が試験にどう反映されているかを調べるような場合では、年度ごとの難易度の違いを考慮しなければならない。当然、受験者集団も異なっているから、難易度の違いを考慮して比較するのは容易ではない。このような場合、難易度の違いは平均点に出ると考えられるが、その違いを反映させる方法はまだ確立されていない。大学入試センター試験の得点調整については、さらにそれを短時間で行わなければならないという制約もあり、そのため分位点差縮小法という方法がつかわれている。この方法は得点調整には向いているが、あくまでも差を縮小するのが目的であるため、異なる難易度の試験結果を同じ土俵の上で分析することには不向きである。

そこで、このような異なる難易度の試験結果をもとに教育効果を検証する方法が求められる。樋口, 伊藤, 三上(2011)[1]では、有限区間の値をとる確率変数を同じ区間上の別の確率変数へと変換するモデルを想定し、一部の学生に実施した講義演習融合型授業の効果について複数年のEMaTの結果をもとに説明した[1]。このような検証は連続した授業の中で行われる複数の試験結果を分析することにも応用が期待できる。Higuchi and Mikami(2011)[2]では、工学技術の継承に関して教育継続に関してのサポートの指針を与えているが、このような研究に対しても、こうした分析手法が役に立つと考えられる。

この論文では、受験者の潜在的な学力と試験の難易度を用いてその試験における正答率を導き出すという考えに基づく得点分布のモデルとその変換について報告する。

2. 得点分布モデル

以下、得点については[0,1]の範囲で考えるものとする。これは得点率と考えれば一般性を失うことはない。受験者のもつ潜在的な学力を[0,1]の値をとる確率変数Xで表す。これは標準的な試験を受けた時の正答率がXであることを想定している。ここである試験に対して正答率がpである受験者が試験を受けたとして、その得点分布について考えてみる。いま試験の問題が同じ配点のn個の小問で構成されているとし、受験者は全ての小問について正答率pで正解するものと仮定する。この仮定のもとでは、受験者の正答数は二項分布$B(n,p)$に従う。すなわち、受験者の得点率は二項分布$B(n,p)$に従う確率変数Xを用いてX/nで表される。一般論を考えるため二項分布を正規分布で近似することにすると、X/nの期待値はp、分散は$p(1-p)/n$であるから、得点率の分布は

*　広島大学大学院工学研究院電気電子システム数理部門

$$p+\sqrt{\frac{p(1-p)}{n}}\varepsilon \tag{1}$$

となる。ただし、$\varepsilon \sim N(0,1)$である。実際にはそれぞれの小問で配点が異なる場合が多いと思われるが、近似的に十分実用があるものと考えられる。

次に潜在的な学力と試験における正答率との関係を考える。潜在的な学力が一定であっても、試験の難易度によって正答率は異なるものと考えられる。ここでは広義の単調増加関数

$$T_\alpha:[0,1]\to[0,1] \tag{2}$$

を考えることにする。したがって、潜在的な学力がxである受験者が難易度αの試験を受けた時の得点は確率変数

$$Y_\alpha=T_\alpha(x)+\sqrt{\frac{T_\alpha(x)(1-T_\alpha(x))}{n}}\varepsilon \tag{3}$$

で表される。さらにある受験者集団の潜在的な学力の分布がXで表されるものとすると、その受験者集団が難易度αの試験を受けた時の得点分布は

$$Y_\alpha=T_\alpha(X)+\sqrt{\frac{T_\alpha(X)(1-T_\alpha(X))}{n}}\varepsilon \tag{4}$$

となる。

次に試験の難易度αとそれに基づく変換T_αについて考える。変換T_αは単調増加関数であれば問題ないと思われるができれば次の条件

$$T_{\alpha\beta}(x)=T_\alpha\left(T_\beta(x)\right) \tag{6}$$

が満たされる方が良いであろう。この条件によって、2つの対象を比較するような場合では、一方を$\alpha=1$と考えてると計算の手間を省くことができる。条件(6)を満たすような変換として、例えば

$$T_\alpha(x)=x^\alpha \tag{7}$$

が考えられる。この場合、難易度αの試験の得点分布は

$$Y_\alpha=X^\alpha+\sqrt{\frac{X^\alpha(1-X^\alpha)}{n}}\varepsilon \tag{8}$$

である。

3. 得点分布の期待値

得点分布(8)がどのような分布であるかは、潜在的学力分布Xに依存する。Xの分布がどのようなものであるかを、調べるため、まずはY_αの期待値と分散を計算してみると次の定理を得る。

定理1：潜在的学力分布をXとするとき、得点分布Y_αの期待値と分散は

$$E(Y_\alpha)=E(T_\alpha(X)) \tag{9}$$

$$V(Y_\alpha)=\frac{n-1}{n}V\left(T_\alpha(X)\right)+\frac{E(T_\alpha(X))(1-E(T_\alpha(X)))}{n} \tag{10}$$

で与えられる。

（証明）　式を簡潔にするため$T_\alpha(X)$をT_αと書くことにする。Xとεの独立性と$\varepsilon \sim N(0,1)$より、

$$\begin{aligned}E(Y_\alpha)&=E\left(T_\alpha+\sqrt{\frac{T_\alpha(1-T_\alpha)}{n}}\varepsilon\right)\\&=E(T_\alpha)+E\left(\sqrt{\frac{T_\alpha(1-T_\alpha)}{n}}\varepsilon\right)\\&=E(T_\alpha)+E\left(\sqrt{\frac{T_\alpha(1-T_\alpha)}{n}}\right)E(\varepsilon)\\&=E\left(T_\alpha(X)\right)\end{aligned} \tag{11}$$

となる。次に分散を計算するため、まずは$E(Y_\alpha^2)$を計算する。

$$\begin{aligned}&E(Y_\alpha^2)\\&=E\left(T_\alpha^2+2\sqrt{\frac{T_\alpha(1-T_\alpha)}{n}}\varepsilon+\frac{T_\alpha(1-T_\alpha)}{n}\varepsilon^2\right)\\&=E(T_\alpha^2)+E\left(2\sqrt{\frac{T_\alpha(1-T_\alpha)}{n}}\right)E(\varepsilon)\\&\qquad+E\left(\frac{T_\alpha(1-T_\alpha)}{n}\right)E(\varepsilon^2)\end{aligned} \tag{12}$$

T_αとεは独立であり、$\varepsilon \sim N(0,1)$より$E(\varepsilon^2)=1$であるから、

$$\begin{aligned}E(Y_\alpha^2)&=E(T_\alpha^2)+E\left(\frac{T_\alpha(1-T_\alpha)}{n}\right)\\&=E(T_\alpha^2)+\frac{E(T_\alpha)-E(T_\alpha^2)}{n}\end{aligned} \tag{13}$$

となる。したがって、

$$\begin{aligned}V(Y_\alpha^2)&=E(Y_\alpha^2)-\{E(Y_\alpha)\}^2\\&=E(T_\alpha^2)-\{E(Y_\alpha)\}^2+\frac{E(T_\alpha)-E(T_\alpha^2)}{n}\\&=V(T_\alpha^2)+\frac{E(T_\alpha)-[V(T_\alpha)+\{E(Y_\alpha)\}^2]}{n}\\&=\frac{n-1}{n}V(T_\alpha^2)+\frac{E(T_\alpha)(1-E(T_\alpha))}{n}\end{aligned} \tag{14}$$

が得られる。　□

したがって、式(7)によって難易度 α に基づく正答率の変換を行う場合、次が成り立つ

系1：潜在的学力分布をXとし、難易度αに基づく正答率変換として$T_\alpha(x) = x^\alpha$を用いるとき、得点分布Y_αの期待値はXの（次数について実数に拡張された）α次モーメント、すなわち

$$E(Y_\alpha) = E(X^\alpha) \tag{15}$$

で与えられる。

4. 難易度の推定

異なる試験の結果に大きく差があった場合、もし受験者集団の分布が同じようなものであるならば、その差は試験の難易度の差によるものと考えられる。定理1より、Y_αの期待値は小問数nに依存しない。そこで期待値を利用した難易度αの点推定を考えることにする。

難易度αに基づく正答率変換として式(7)を用いるとすると、系1よりY_αの期待値はXのα次モーメントとなる。しかし、通常Xについては未知である。そこで一方の試験結果の難易度を1とすることによりY_1とX^α の関係から、$E(X^\alpha)$を導出する。

まず、x^α を漸近展開すると

$$x^\alpha = x_0^\alpha + \alpha x_0^{\alpha-1}(x - x_0) + O((x - x_0)^2) \tag{16}$$

であるから、

$$\begin{aligned} Y_1^\alpha &= \left(X + \sqrt{\frac{X(1-X)}{n}}\varepsilon \right)^\alpha \\ &= X^\alpha + \alpha X^{\alpha-1}\left(\sqrt{\frac{X(1-X)}{n}}\varepsilon \right) \\ &\qquad + O\left(\left(\sqrt{\frac{X(1-X)}{n}}\varepsilon \right)^2 \right) \end{aligned} \tag{17}$$

となる。ここで両辺の期待値をとると

$$\begin{aligned} E(Y_1^\alpha) &= E(X^\alpha) + E\left[O\left(\left(\sqrt{\frac{X(1-X)}{n}}\varepsilon \right)^2 \right) \right] \\ &\approx E(X^\alpha) \end{aligned} \tag{18}$$

を得るから、

$$E(Y_\alpha) = E(X^\alpha) \approx E(Y_1^\alpha) \tag{19}$$

に基づくαの点推定を考えればよい。

同一の受験者集団とみなせる２つの試験A, Bの得点に関するデータとして、$\{a_i\}(i=1,\cdots,m)$, $\{b_j\}(j=1,\cdots,n)$があるとする。試験Aの難易度を1としたとき、試験Bの難易度αについての点推定として、以下の方法を提起する。まず難易度Aの試験データをもとに経験 α 次モーメント

$$M(\alpha) = \sum_i a_i^\alpha \tag{20}$$

を計算し、$E(Y_1^\alpha)$の近似値とみなす。$M(\alpha)$はαの関数であるから、試験Bにおける平均値と一致する$\hat{\alpha}$をαの推定値とすればよい。$M(\alpha)$は単調減少関数で

$$\lim_{\alpha\to 0} M(\alpha) = 1, \lim_{\alpha\to\infty} M(\alpha) = 0 \tag{21}$$

であるから、推定値 $\hat{\alpha}$ は必ず存在し、一意に定まる。人工的なデータをもとに$M(\alpha)$を計算した結果をFig. 1に示す。

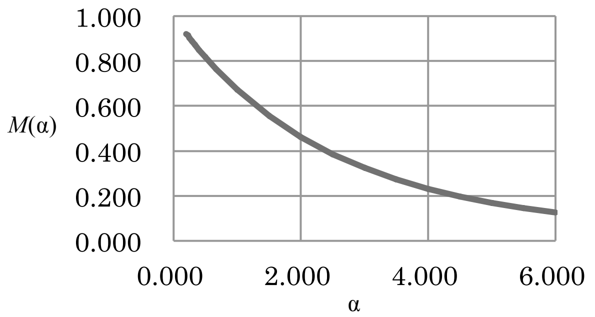

Fig.1 The graph of $M(\alpha)$

5. 考察

樋口，伊藤，三上(2011)[1)]ではこの方法で難易度の推定を行い、講義演習融合型授業のグループに教育効果があることを示した。授業改善による教育効果を示すためには、同程度とみられる学生をグループ分けして別の授業を行ったうえで、難易度が同程度の試験をグループごとに行うなどすればよい。しかし、同じ課程に属する学生を教育効果が異なると思われる授業にグループ分けすることは、教育組織として問題がある。特定の課程に対して実験的に授業を行った場合、同じ難易度の試験を実施することが難しくなる。そのため、授業方法を変えない課程の試験結果をもとに難易度の違いを推定することで、授業方法を変えた課程について難易度を考慮したうえで改善を確認することができるはずである。そのためにこのような推定法を確立させる必要がある。

現時点でこの方法が抱える問題点として、まずはテストの小問についての独立性が考えられる。誘導形式の問題の場合、ある問題で正答できなかった受験者は次の問題も正答できない可能性が高く、小問ごとに独立とはい

えない。しかし、小問ごとに独立と考えられるような試験も確実に存在するので、少なくともそのようなテストに関しては、問題がないといえる。次に小問数 n に関する問題がある。小問数 n に関する項が残っていることからわかるように、n が十分大きくないと推定の誤差が大きくなると考えられる。また小問の数え方の問題が出てくるので、n の影響ができるだけ少なくなるモデルを考えた方がよいであろう。さらにこのモデルでは、期待値の計算は比較的易しいといえるが、分散の計算は容易ではない。分散の計算になるとさらに小問数 n の影響が大きくなると予想されるため、この点を解決できるモデルを作る必要がある。

また、このモデルが現実的かどうかの考察がまだ進んでない。そのためには、数値実験で得点の分布がどのように変換されるかを調べ、実際の試験における得点分布を精査する必要がある。例えば$X\sim B(20,0.5)$とした場合にいくつかのαについてY_αの分布を調べるとFig. 2,3,4のようになる。さらにこのモデルにあてはまる分布を調べていくことで受験者集団の潜在的学力の分布についても踏み込んだ研究が展開できると期待できる。

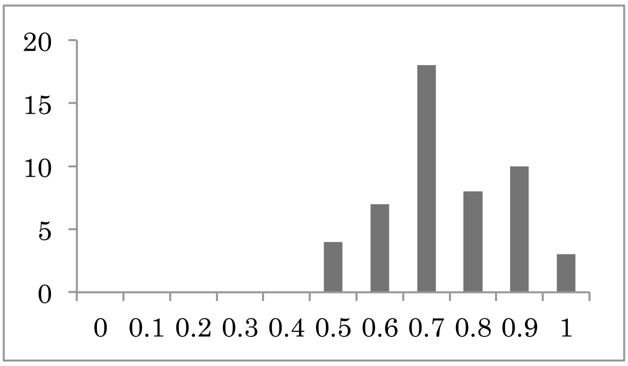

Fig.2 The distribution of $Y_{0.5}$

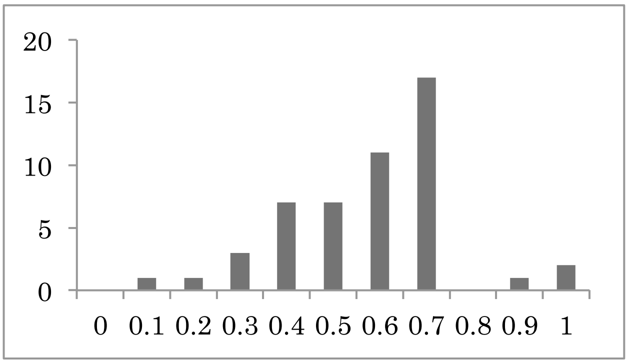

Fig.3 The distribution of Y_1

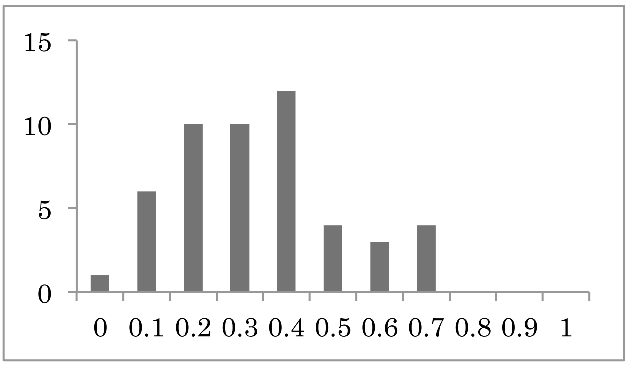

Fig.4 The distribution of Y_2

6. まとめ

この論文では潜在的な学力と試験の難易度からその試験に対する正答率が導かれるというモデルを作成し、そのモデルの下で、期待値と分散を計算し、難易度の推定法を提起した。試験の得点分布には不確定な要素が多いとともに、問題作成が人為的なものであるため、深くデータを分析していくことは難しい。今回作成したモデルでは、大胆な仮定をいくつか導入しているが、その仮定が妥当かどうかは今後の研究で明らかとなるだろう。

このモデルによる分析により、授業の途中で実施された試験結果から、その授業の教育効果を確認することができるようになる可能性がある。Higuchi and Mikami(2011)[2)]では工学技術の継承における人的サポートの指針を示しているが、一般の教育においても、その課程における試験成績の変化を難易度の違いをできるだけ排除した上で分析し、モチベーションだけではなく成績を踏まえたうえで、教育継続に対するサポートの方向性を示すことが期待できるはずである。

参考文献

1) 樋口勇夫，伊藤浩行，三上敏夫：数学基礎教育における講義演習融合授業の効果についてー教育GP「工学教育を支える「数学力」養成プログラム」の取組としてー，平成23年度工学教育研究講演会講演論文集，pp.610-611，2011.

2) Higuchi, I. and Mikami, T., Maxima and minima of overall survival functions with fixed marginal distributions and transmission of technology, *Communications in statistics –Theory and Methods*, 41, pp. 46-61, 2011.

平成23年10月28日　受理