

CBTにおける受検者の回答速度特性推定の試み

藤田彩子 舩田博之

株式会社 リクルートマネジメントソリューションズ 組織行動研究所

背景と目的

CBT が普及していく中で、単に回答の正誤だけでなく、さまざまな情報を得ることができるようになっている。中でも回答時間については、項目の特性として捉えれば、課題の複雑さ、煩雑さを表す指標となり、受検者の個人特性として捉えれば情報処理・判断のスピードを表す指標となるなど、有用な情報である。こうした有用性から回答時間に関する研究も進められており、モデル化についても、対数正規分布を用いたモデル (van der Linden, et al., 1999, 以下 Lognormal モデルとする)、 χ^2 分布を用いて回答時間分布を再現するモデル(永岡 他 1989, 以下 分布モデルとする)、IRT の考え方を応用したモデル(永岡 他 1991)など、いくつかあり、それぞれに利点や課題点がある。

回答時間は項目と受検者の双方に依存して決まるため、IRT のように規準集団という軸を定めた上で、項目特性や受検者の個人の回答速度特性をパラメータ化する必要がある。また、実際の場面で活用していく場合には、一定の安定性を持った指標であることはもちろんであるが、その指標の算出が容易で扱いやすいことが求められる。この点、 χ^2 分布を応用する手法は、パラメータの算出が簡単で扱いやすい。また、パラメータの解釈も容易で、項目の特徴を回答時間の観点から捉えることができる。ただし、 χ^2 分布自身は単に回答時間分布を予測する分布であるというだけで、受検者の回答速度特性を規準集団内で位置づけるには、何らかの工夫が必要である。そこで、本研究では、 χ^2 分布モデルの回答時間分布モデルとしての有効性を Lognormal モデルとの比較によって確認し、受検者個人の回答速度特性の推定を試みることを目的とする。

分布モデル

分布モデルは、回答所要時間データの発生モデルをエントロピー最大化原理により理論的に導出したもので、パラメータの推定と解釈が非常に容易であることが特徴である(植野 2003)。

分布の密度関数は、

$$f(t) = \frac{t^{a-1} \exp(-t/b)}{b^a \Gamma(a)}$$

で示され、回答所要時間の累積分布 $F(t)$ は、

$$F(t) = \begin{cases} 0 & t < t_0 \\ \int_{t_0}^t f(t) dt & t \geq t_0 \end{cases}$$

で示される。ここで、 t_0 は、各項目における、回答に必要な最小限の時間を表す。 a 、 b は、それぞれ

$$\begin{aligned} a &= E^2/V \\ b &= E/a = V/E \end{aligned} \quad (E: \text{実データの平均から } t_0 \text{ をひいたもの}, V: \text{実データの分散})$$

である。 と の意味は、 は問題に回答するのに必要な「読む」「考える」などの単一要素過程の数（複合度）（植野 2003）、 は単一要素過程あたりに要する時間と解釈される。

Lognormal モデル

Lognormal モデルは、受検者 j が項目 i に回答したときの回答時間 T_{ij} を

$$\ln T_{ij} = m + d_i + t_j + e_{ij}$$

ただし、

μ : すべての受検者のすべての回答項目についての回答時間の自然対数の平均で $m \equiv E_{ij}(\ln T_{ij})$

i : 項目 i についての回答時間の自然対数の平均で、 $d_i \equiv E_j(\ln T_{ij}) - m$

j : 受検者 j についての回答時間の自然対数の平均で、 $t_j \equiv E_i(\ln T_{ij}) - m$

e_{ij} : 誤差項で正規分布 $N(0, \sigma^2)$ に従う。 $s^2 = E_{ij}(\ln T_{ij} - d_i - t_j - m)^2$

で表すことができる、というモデルである。受検者の回答時間特性と項目の回答時間特性を独立とおき、この二つの特性それぞれの、項目プール・受検者全体の回答時間平均からの相対的位置によって T_{ij} を表そうとするものである。ゆえに、パラメータは項目プール・受検者の双方に依存し、項目プールに新しい項目を追加して μ が変わってしまうような場合には、既存の項目のパラメータも算出しなおす必要がある。

手続き

用いたデータ

株式会社リクルートマネジメントソリューションズが開発した基礎能力検査の下位検査である、非言語検査(数量的処理や論理の問題項目からなる)に対する 8283 人の回答所要時間データを用いた。28 題(どの問題も 1 題 2 問組となっている)の項目プールから 3 題を組にした項目セットが 15 あり、受検者にはこの 15 セットの中からランダムに一つが割り当てられる。1 題ずつの回答時間が記録されるため、一人につき 3 題分の回答時間データが得られた。

このデータを用いて 分布のパラメータ を求め、回答時間分布を再現した。次に分布の累積分布関数を用いて、得られた受検者の項目ごとの回答時間が、集団の中でどの水準にあるのかを算出し、受検者の回答時間特性は正規分布に従うという仮定のもと、その累積パーセントを標準正規分布に置き換えた。この操作によって、回答時間分布が項目ごとに異なっても、同じ指標で受検者の回答時間を評価できるようになる。3 題それぞれの z -score を「項目ごとの回答速度スコア」ととらえ、3 題の回答速度スコアの平均値をその受検者の「回答速度」とした。

同じデータに対して Lognormal モデルを適用し、 T_{ij} と e_{ij} の関係からこのモデルにおける回答時間分布の再現性を確認し、受検者の「回答速度」を求めた。

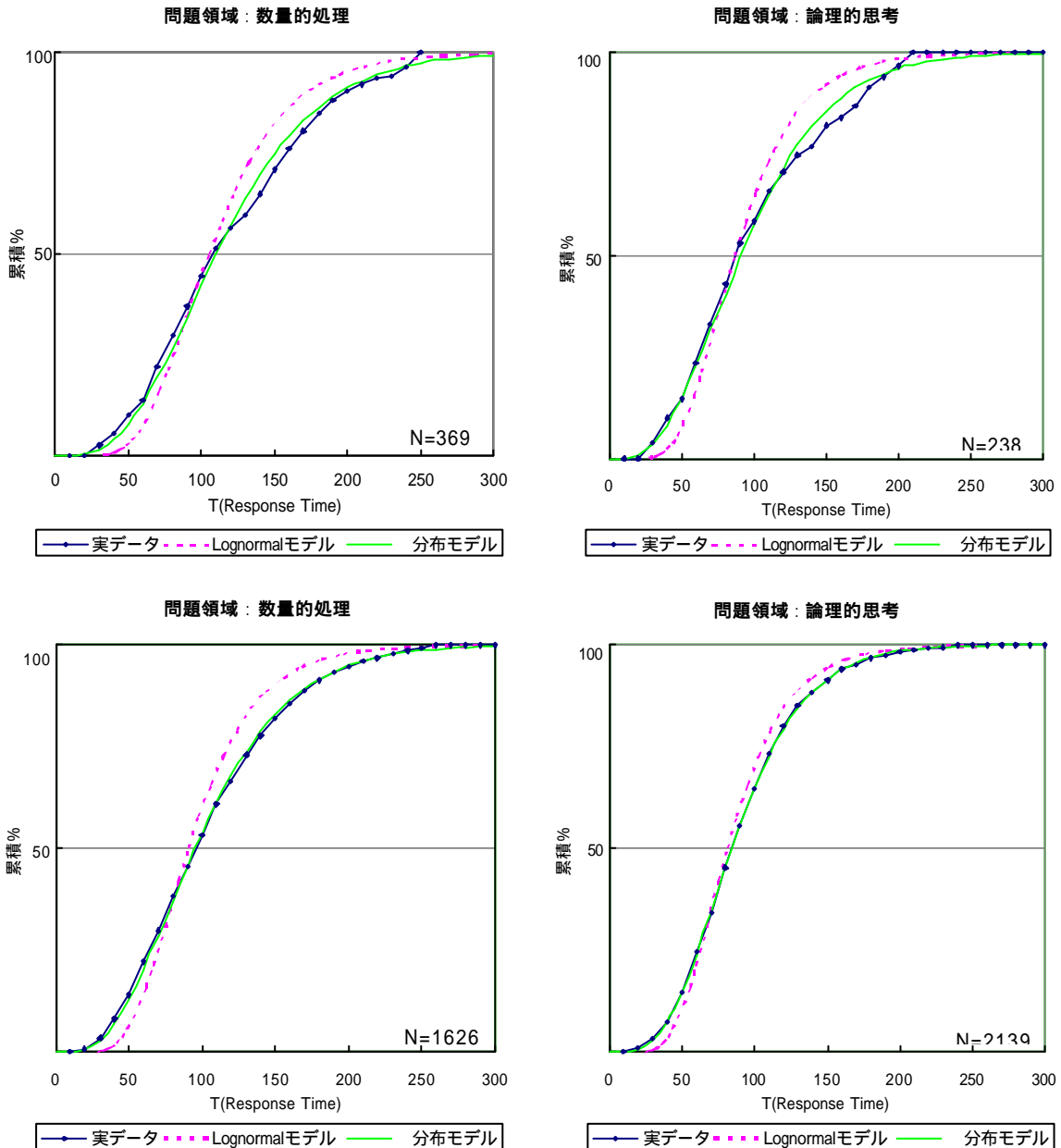
結果

分布モデルと Lognormal モデルによる回答時間分布の再現性の比較

分布モデルは、受検者集団の回答時間分布を非常によく再現した。データが 300 に満たないような場合でもかなりの精度で実際の累積分布を再現でき、安定性の高いモデルであることが確認された。Lognormal モデルでは、項目によっては実際の回答時間分布をかなりよく再現できる場合もあったが、すべての項目について 分布モデルよりも、再現性が悪かった。また、数量的処理が中心となる問題領域では再現性が全体的に悪く、論理的な思考を問う問題領域では再現性がよいなど、領域によって再現性に違いがあった(図 1)。

Lognormal モデルのほうがモデルとしては単純であるため、データが少ない場合や項目に何らかの特徴があるなど、回答時間分布の形がいびつになるような場合は、うまく分布を再現できないという可能性がある。

図1 分布・Lognormal モデルの実データに対するあてはまり



個人の回答速度特性の推定値の信頼性

分布モデルから求めた受検者の「回答速度」の信頼性を確認するため、係数を算出したところ、15 セットの 係数の平均値は 0.563 で、標準偏差は 0.043 であった。3 題のデータでの数値であることを考えると、十分高い水準であるといえる。スピアマン・ブラウンの公式によれば、項目数と信頼性との関係から、9~10 題のデータがあれば、係数は 0.8 程度になる。分布モデルから求めた「回答速度」を Lognormal モデルから求めた「回答速度」と比較したところ、相関は 0.99 で、ほぼ同じ結果となることが確認された(図 2)。また、どちらも推定された回答速度特性は正規分布に従うことが確認された(図 3)。

図2 分布モデルとLognormal モデルでの
受検者の回答速度特性の一致度

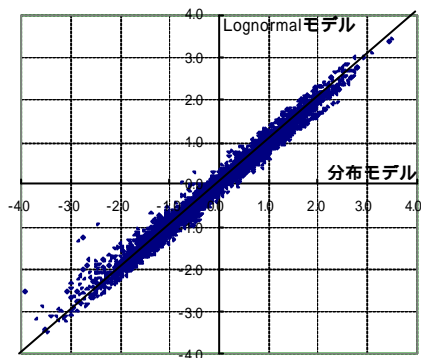
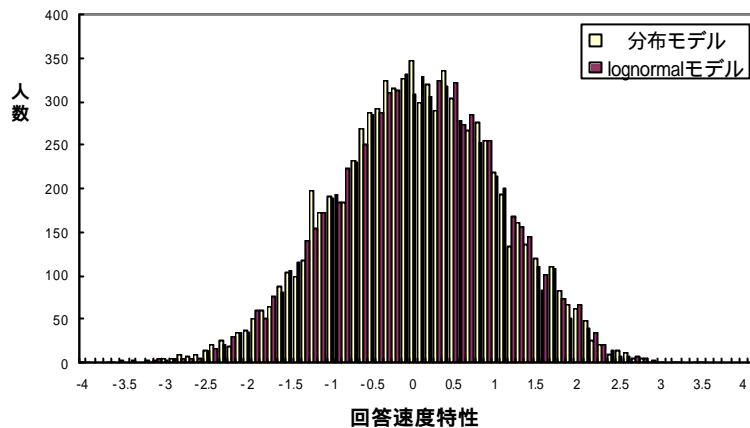


図3 分布モデルとLognormal モデルでの回答速度特性の分布



まとめ

今回の研究から、分布モデルが回答時間分布を非常によく再現できること、またデータが比較的少なくても安定的にパラメータが得られることが分かった。しかもパラメータの推定は容易であり、非常に実用的なモデルといえる。また、受検者の項目ごとの回答時間を分布の中に位置づけ、正規化することにより、受検者個人の回答速度特性をパラメータ化することができた。Lognormal モデルを用いた場合も、回答時間分布を簡単に再現できるが、項目によっては当てはまりの悪い場合もあった。

このようなモデルを利用した回答速度特性と能力特性とあわせることにより、より詳細かつ多面的に個人の特徴を捉えることができる。このことは、e-learning などの学習用コンテンツにおいても、正誤だけでなく速度も含めて学習者の理解度を推定するなど、さまざまな応用が考えられる。また、項目の特性としてみれば、適切なテスト時間・学習時間の見積もりや時間特性の異なる項目の排除など、テスト設計・学習用コンテンツ設計においてもさまざまな活用が期待できる。

今後は受検者の項目に対する正誤と回答時間の関係、問題領域による回答時間の違いなどを確認し、より高い精度での回答速度特性を求めるための改善点を探っていく必要がある。また、分布モデルの定数や個人の回答速度特性など、パラメータ算出方法についても、改良の余地を検討する必要がある。

参考文献

- van der Linden, W.J., Scrams, D.J., & Schnipke, D.L. (1999). Using response-time constraints to control for differential speededness in computerized adaptive testing. *Applied Psychological Measurement*, 23, 195-210
- 永岡慶三 植野真臣(1991) 回答所要時間における項目応答理論 行動計量学 18巻2号,1-8
- 永岡慶三 吳亜棟(1989) コンピュータ・テストングにおける回答所要時間についての分析 日本教育工学雑誌 12(4) 129-137
- 植野真臣(2003) ガンマ分布によるWBT 所要時間データのオンライン解析システム 教育システム情報学会 第28回全国大会論文集
- 植野真臣(2004) eラーニングにおける学習履歴データのデータ/テキスト・マイニング 日本行動計量学会第7回セミナー講演論文集