

オンライン教材におけるコンバージョンの多重ハードルモデル

草 薙 邦 広

広島大学外国語教育研究センター

1. はじめに

本稿は、オンライン英語学習教材における学生のコンバージョン (conversion, CV) を説明する数理モデルを探索的に検討するものである。オンライン英語学習教材に対するログイン数の経験的分布は、計数データに対して慣習的に近似されるポアソン分布や負の二項分布によって適切に近似することができない場合が多い。それと同時に、コンバージョンの意思決定過程といった観点からは、ポアソン分布や負の二項分布による近似は、データ発生機序としてのもっともらしさを欠く。一方、コンバージョンに至る過程は多段階を経るという想定に対して整合的な、本稿が提案するところのハードル負の二項分布および負の二項分布からなる混合分布 (二重ハードルモデル) は、データ発生機序の観点からも、より妥当なモデルだと予想される。本稿は、当該の検証モデルをマルコフ連鎖モンテカルロ法によるベイズ推定によって観測に近似させたところ、結果として、このモデルが比較的優れた近似を示したという事例を報告する。

コンバージョンとは、情報通信分野、特にその応用である e コマースや Web マーケティング分野において活用される概念であり、Web サイト等において期待される一種の効用や成果を参照する。通例的に、e コマースサイトでは商品の購入、SNS では会員登録、広告サイトでは問い合わせといった事象をコンバージョンとみなす。商品購入の例では、Web サイトに訪れたユーザ数に対して商品を購入したユーザ数の比は、コンバージョン率 (conversion rate, CVR) と呼ばれ、実務上、この値の最大化が図られる (e.g., McDowell, Wilson, & Kile, 2016)。なお、コンバージョン率の最大化を目標とした一連の施策は、コンバージョン最適化 (conversion optimization) と呼ばれる。コンバージョンは、外国語教育研究やコンピューター支援語学学習 (CALL) において広く知られる学術用語ではないが、教育における質保証の観点や、データ駆動主義が重要視される近年の情勢下において有益な概念になると考えられる。

このコンバージョンという概念をオンライン英語学習教材に対して援用するならば、コンバージョンを標準テストで測定される学習成果や、履修科目成績に据えるべきであると考えられる。一方で、必ずしもオンライン英語学習教材における利活用の度合いが学習成果の直接的な原因になるとは限らないという報告もある (cf. Henrie, Halverson, & Graham, 2015; Li & Tsai, 2017)。また、学生が自律的で継続的な学習習慣を獲得すること自体も、長期的な視点に経てば重要な教育目標であるともいえる。このことから、教材に対するログイン回数、ログイン時間、消化コンテンツ数といった学習履歴それ自体もまた、一種の重要なコンバージョンであるとみなすこともできるはずである。

コンバージョンの分析では、任意のコンバージョンに対して、マイクロコンバージョンと呼ばれる下位過程を想定することが多い。たとえば、e コマースを例とすると、ユーザは商品検索の反復によって商品を選択し、その後決済手続きを取る。この場合では、決済手続きがコンバージョンであり、商品検索の過程がマイクロコンバージョンである。

オンライン英語学習教材に適用すると、まず、当該教材に対して 1 度ログインする過程がある

と想定する。本稿はこれをマイクロコンバージョン1と呼ぶ。その後、ユーザはログインを反復し、定期的にログインを反復する習慣を身につけると想定する。この学習習慣の獲得をマイクロコンバージョン2とする。その後、ログインの反復に伴って、規定のコンテンツ数すべてを消化した状態に至るはずである。これを最終的なコンバージョンとする。これに加えて、履修登録をした学生が、マイクロコンバージョン1に遷移する過程をハードル1、マイクロコンバージョン1からマイクロコンバージョン2への過程をハードル2とする。最後に、マイクロコンバージョン2からコンバージョンへの過程がハードル3である。これらの関係を状態遷移図の要領によって可視化したものが図1である。

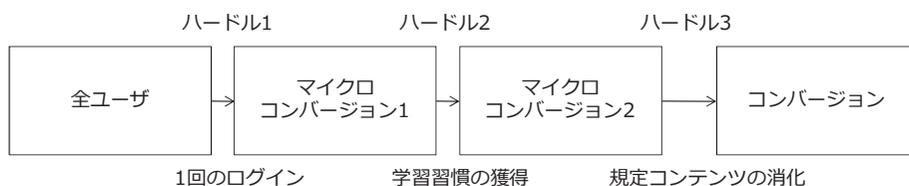


図1 本稿におけるコンバージョンとハードルのコンセプト

2. 当該データ

コンバージョンとハードルによるコンセプトの有効性を検証するために、本稿は広島大学外国語教育研究センターが提供する、2019年度第3ターム開講科目「コミュニケーションIIA」において使用されるオンライン学習教材「ぎゅっとe」の学習履歴データを使用した。当該の教材において、履修登録があった909名の第3ターム終了時におけるログイン回数および消化コンテンツ数のみを分析対象とした。そのうち、10名のデータはログが適切に記録されていないため、分析から除外した。当該教材のコンテンツは、授業計画において授業回のテストと連動しており、ターム終了時点までに、560チャプターの消化が学生に課される計算になる。

図2は、ログイン回数の経験度数分布である。一見して複雑な分布形状を示していることが分かる。特にログイン数0回が突出している傾向が見て取れる。なお、このデータの期待値は16.70、分散は177.56である。表1は、ログイン回数の分位数による要約であり、図3は経験累積分布関数(ECDF)を示している。図4はローレンツ曲線である。ローレンツ曲線の視察から、ユーザ間において格差が比較的大きい傾向が把握できる。たとえば、全ログイン数のうち半数は、上位25%程度のユーザが占めている。

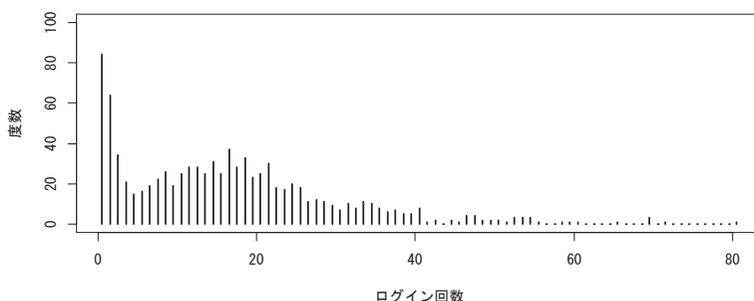


図2 ログイン回数データの度数分布 (n = 899)

表1 ログイン回数データの要約

分位数	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
値	0	2	3	8	12	15	18	22	26	34	81

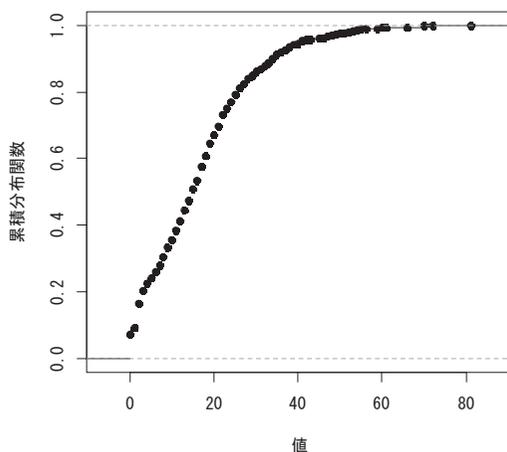


図3 ログイン回数データの経験累積分布関数

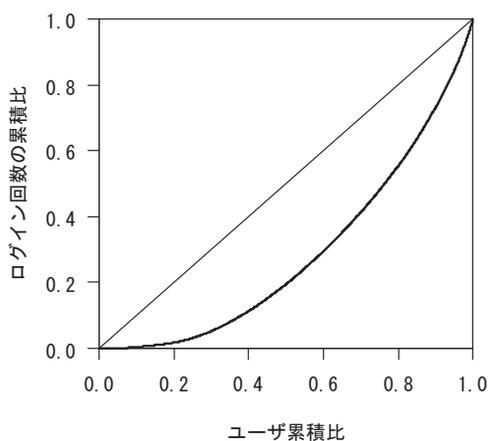


図4 ログイン回数データのローレンツ曲線

なお、909名すべてのユーザのうち、履修を取りやめた学生や、履修を続けてはいるが、ターム内においてライセンス料を支払わずに、やむを得ず当該教材について学習停止となった学生も含まれている。一方で、ターム内において学習停止になったからといって、当該の教材を利用しないわけではない。表2は、ログインの有無、学習継続および学習停止によってユーザ数をクロス集計したものである。なお、本稿では、ターム内において学習停止になったユーザのデータも分析に含めてある。ログインのなかったユーザ数は、計65名であり、これはデータ分析対象ユーザの7.23%に相当する。

表2 ログインの有無と学習継続および学習停止のクロス集計 ($n = 899$)

	ログインあり	ログインなし	計
学習継続	805	15	820
学習停止	29	50	79
計	834	65	899

次に、消化コンテンツ数のデータについて検討する。図5が示している消化コンテンツ数の経験度数分布は、一見して二分化された傾向を示している。参考として、期待値は363.02であり、分散は54986.43である。期待値である363.02は、授業計画上で目標とされている消化コンテンツ数560に対して、およそ65%に相当する。

本稿は、560チャプター以上のコンテンツ消化を最終的なコンバージョンとみなし、コンバージョンしたかどうかを2値データのダミー変数として扱い、CVと記す。コンバージョンした場合は $CV = 1$ であり、コンバージョンしなかった場合の値は0である。CV = 1のユーザは、393名(.44)であり、CV = 0のユーザは506名(.56)であった。つまり、CVRは.44である。

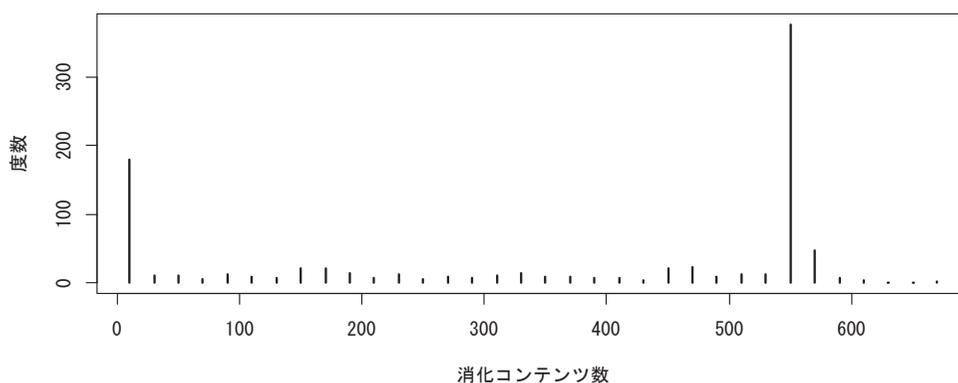


図5 消化コンテンツ数の度数分布

3. モデリング

3.1 方針

本稿のモデリング過程は2つのステップから構成される。最初のステップは、コンバージョンとハードルによるモデルを検証するために、上記のログイン回数の度数分布に近似する確率分布を検討することである。ログイン回数の度数分布に近似するモデルとして、6つの比較モデルと1つの検証モデルをマルコフ連鎖モンテカルロ法によるベイズ推定によって近似させ、それらの適合度を比較する手続きを取る。本稿は、検証モデルが比較モデルよりも優れた適合度と予測精度を示すことを期待している。

本稿における検証モデルの大まかな自然言語上の含意は、以下の通りである。第一に、全ユーザのうちハードル1を超えないユーザの比率があり、この比率に特徴づけられるユーザのログイン回数は確率1で0となる。次に、ハードル2を超えないユーザの比率があり、この比率に特徴づけられるユーザに対応するログイン回数が従う確率分布を推定する。最後に、ハードル2を超えたユーザの比率と、このハードル2を超えたユーザのログイン回数が従う確率分布を推定する。

より具体的に、検証モデルが想定するデータ発生機序は、以下の通りである¹⁾。ただし、想定するデータの発生機序と、検証モデルにおける推定の統計的手順が同一であるとは限らない。

- (a) ベルヌーイ分布によって、ハードル1を超えるかどうか決定される
- (b) ハードル1を超えないユーザのログイン回数は確率1をもって0となる
- (c) ユーザがハードル2を超えたか、ユーザがハードル2を超えなかったかが一定の比率で決定される
- (d) ハードル1を超え、かつハードル2を超えないユーザのログイン回数は、切断負の二項分布によって確率的に生成される
- (e) ハードル2を超えたユーザのログイン回数は、負の二項分布によって確率的に生成される

この検証モデルを、本稿は多重ハードルモデル²⁾と呼ぶ。以下に紹介するように、多重ハードルモデルの実装の1つとして、ハードル負の二項分布と負の二項分布による有限混合分布モデルを想定した。

次のステップでは、最終的なコンバージョンを示すダミー変数CVは、上記多重ハードルモデルによって近似されるログイン回数を項とする関数とみなせることを示す。つまり、ログイン回数が増えることにより、コンバージョンとなる確率が増加していく傾向を明らかにする。

3.2 数理的基盤

検証モデルの定式化の前に、まずはモデリングにおいて使用する数理的道具立てについて解説する³⁾。まず、個人*i*のログイン回数を $k(k = 1, 2, 3\dots)$ とし、 k の確率質量関数 $f(k)$ について考えることとする。ハードル1を超えるユーザと、ハードル1を超えないユーザがいるということに着目して、 $f(k)$ において $k = 0$ という事象と $k > 0$ という事象がともにベルヌーイ分布に従うと仮定すると、

$$p(k = 0) = \pi \tag{1}$$

$$p(k > 0) = 1 - \pi \tag{2}$$

が得られる。この確率 π は、定義通り、ハードル1を超えない確率である。

次に、 $k > 0$ の条件下における k の確率分布として、ポアソン分布を仮定することにする。ポアソン分布は、計数データに対して最も頻繁に使用される離散確率分布の1つであり、母数 λ のみをもつ。ポアソン分布の確率質量関数は、

$$f(k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{3}$$

であるが、この標準的なポアソン分布は $k = 0$ を取る確率をもつ。そのため、比率 π によって0を取る確率と、ポアソン分布によって0を取る確率が重複することにより、確率質量関数の積分が1を超える場合もあり得る。そのため、確率質量関数の台を正の整数のみに限定した新たな確率分布を使用しなければならない。確率質量関数の積分が1となるように調節すると、 $k = 0$ の確率が0、かつ $k > 0$ の条件下における k の確率質量関数 $g(k)$ は、

$$g(k) = \begin{cases} 0 & \text{if } k = 0 \\ \frac{f(k)}{(1-f(0))} & \text{if } k > 0 \end{cases} \quad (4)$$

となり，これはゼロ切断分布（zero-truncated distribution）と呼ばれる。この式の関数 f に対して，標準的なポアソン分布の確率質量関数を代入すると，

$$g(K = k | k > 0) = \begin{cases} 0 & \text{if } k = 0 \\ \frac{\lambda^k}{(e^\lambda - 1)k!} & \text{if } k > 0 \end{cases} \quad (5)$$

となり，これがゼロ切断ポアソン分布の確率質量関数である。よって，ベルヌーイ施行によって0または0以上の値を取り，次に0以上の値を取る場合に限って，正の整数のみを台にもつゼロ切断ポアソン分布に従う値を取る確率分布の確率質量関数は，

$$p(k|\pi, \lambda) = \begin{cases} \pi & k = 0 \\ (1-\pi) \frac{\lambda^k}{(e^\lambda - 1)k!} & k > 0 \end{cases} \quad (6)$$

となる。これを，ハードルポアソンモデルという。確率分布として強調する際には，ハードルポアソン分布と呼ぶ。このハードルモデルは，ポアソン分布に限らず，同様に計数データに対して使用される負の二項分布に対しても応用することができる。負の二項分布の確率質量関数は，母数 r と p を使用して，

$$f(k|r, p) = \binom{k+r-1}{k} p^r (1-p)^{k-r} \quad (7)$$

である。ここでは，二項係数を使用している。しかし，分析上の利便性から，本稿では，

$$f(k|\mu, \theta) = \binom{k+\theta-1}{k} \left(\frac{\mu}{\mu+\theta}\right)^k \left(\frac{\theta}{\mu+\theta}\right)^\theta \quad (8)$$

という母数化（parameterization）を使用した。この負の二項分布を使用して，ポアソン分布の場合と同様に，ハードル負の二項モデルとハードル負の二項分布も定義することができる。

ハードルモデルは，値0の事象について，ベルヌーイ施行のみによる0の確率を考慮するが，先に述べたように，ポアソン分布や負の二項分布の母数次第では，ベルヌーイ施行とは独立して，0の値が発生する確率が0以上となる場合もある。この0以上の確率を同時に計算するモデルがゼロ過剰モデル（zero-inflated model）である。たとえば，ゼロ過剰ポアソン分布の確率質量関数は，

$$p(k) = \begin{cases} \pi + (1-\pi)e^{-\lambda} & k = 0 \\ (1-\pi) \frac{\lambda^k e^{-\lambda}}{k!} & k > 0 \end{cases} \quad (9)$$

である。 $k=0$ のとき，標準的なポアソン分布の確率質量関数に0を代入すると，

$$p(0|\lambda = 0) = e^{-\lambda} \quad (10)$$

となる。つまり、(9)式における $k = 0$ の確率は、ベルヌーイ分布によって 0 を取る確率と標準的なポアソン分布によって 0 を取る確率の和である。ゼロ過剰ポアソン分布と同様に、ゼロ過剰負の二項分布も定義できる。

次に、有限混合分布モデル (finite mixture model) について考える。確率変数 $X(x_i = x_1, x_2, x_3 \dots x_N)$ に関する混合分布モデルにおける一般式 $g(x_i)$ は、混合する分布の数を $J(j = 1, 2, 3 \dots J)$ 、 f_j をそれぞれの確率密度分布ないし確率質量分布とすると、

$$g(x_i) = \sum_{j=1}^J \pi_j f_j(x_i) \quad (11)$$

である。ここでの π は混合比率母数である。

さて、これまで取り上げたハードルポアソン分布、ハードル負の二項分布、ゼロ過剰ポアソン分布、そしてゼロ過剰負の二項分布などを混成した混合分布を作成することも可能である。これによって、本稿の検証モデルである、ハードル負の二項分布と負の二項分布による混合分布を表現することができる。

3.3 ベイジアンモデリング

本稿では、最初のステップにおける比較モデルとして、以下の 6 つを想定した：(a) ポアソン分布、(b) 負の二項分布、(c) ハードルポアソン分布、(d) ハードル負の二項分布、(e) ゼロ過剰ポアソン分布、そして (f) ゼロ過剰モデル負の二項分布である。これらの母数を、マルコフ連鎖モンテカルロ法 (MCMC) を使用してベイズ推定する。ベイズ推定には、計算環境 R (R Core Team, 2016) およびそのパッケージである brms (Buerkner, 2017) パッケージ、そして統計モデリング言語であり、ハミルトニアン・モンテカルロ法によるサンプラーである Stan (Carpenter, et al., 2017; Stan Development Team, 2018) を使用した。以降、確率的プログラミング言語に沿うように各モデルを定式化する。

モデル 1 ; $k_i \sim \text{Poisson}(\lambda)$ (12)

$$\lambda \sim \text{Student_t}(\nu = 3, \mu = 3, \sigma = 10) \quad (13)$$

モデル 1 のポアソン分布は、個人 i についてのログイン回数 k_i は、母数 λ のポアソン分布によって発生するというモデルである。本稿において、(12) 式の「 \sim 」記号は、左辺が右辺の確率分布に従う、または確率質量関数であることを示す。(13) 式は、 λ の事前分布を表している。ただし、本稿における事前分布は、すべて十分に広い値を取るよう与えたものである。

モデル2 : $k_i \sim \text{NegativeBinomial}(\mu, \theta)$ (14)

$$\mu \sim \text{Student_t}(\nu = 3, \mu = 3, \sigma = 10) \quad (15)$$

$$\theta \sim \text{Gamma}(\alpha = \frac{1}{100}, \beta = \frac{1}{100}) \quad (16)$$

モデル2の負の二項分布は、ログイン回数が負の二項分布に従うことを示している。負の二項分布の母数は μ と σ の2つである。

モデル3 :

$$p(x) = \begin{cases} \pi & \text{if } k = 0 \\ (1 - \pi) \frac{\text{Poisson}(k|\lambda)}{1 - \text{PoissonCDF}(k|\lambda)} & \text{if } k > 0 \end{cases} \quad (17)$$

$$\lambda \sim \text{Student_t}(\nu = 3, \mu = 3, \sigma = 10) \quad (18)$$

$$\pi \sim \text{Logistic}(\mu = 0, \sigma = 1) \quad (19)$$

モデル3のハードルポアソン分布は、ログイン回数がハードルポアソン分布に従うことを示している。この分布の母数はポアソン分布に π が加えられる。 π は、 $k = 0$ の確率を示す。

モデル4 :

$$p(x) = \begin{cases} \pi & \text{if } k = 0 \\ (1 - \pi) \frac{\text{NegativeBinomial}(k|\mu, \theta)}{1 - \text{NegativeBinomialCDF}(k|\mu, \theta)} & \text{if } k > 0 \end{cases} \quad (20)$$

$$\mu \sim \text{Student_t}(\nu = 3, \mu = 3, \sigma = 10) \quad (21)$$

$$\theta \sim \text{Gamma}(\alpha = \frac{1}{100}, \beta = \frac{1}{100}) \quad (22)$$

$$\pi \sim \text{Logistic}(\mu = 0, \sigma = 1) \quad (23)$$

同様に、モデル4のハードル負の二項分布の分布の母数は π を加えて3つである。

モデル5 :

$$p(k_i) = \begin{cases} \pi + (1 - \pi) \times \text{Poisson}(0|\lambda) & \text{if } k_i = 0 \\ (1 - \pi) \times \text{Poisson}(0|\lambda) & \text{if } k_i > 0 \end{cases} \quad (24)$$

$$\lambda \sim \text{Student_t}(\nu = 3, \mu = 3, \sigma = 10) \quad (25)$$

$$\pi \sim \text{Beta}(\alpha = 1, \beta = 1) \quad (26)$$

モデル5はゼロ過剰ポアソン分布である。ゼロ過剰ポアソン分布の母数は3つであるが、ハードルポアソン分布とは異なり、 π の推定値がそのまま0を取る確率だとは限らない。

モデル6 :

$$p(k_i) = \begin{cases} \pi + (1 - \pi) \times \text{NegativeBinomial}(0|\mu, \theta) & \text{if } k_i = 0 \\ (1 - \pi) \times \text{NegativeBinomial}(0|\mu, \theta) & \text{if } k_i > 0 \end{cases} \quad (27)$$

$$\mu \sim \text{Student_t}(\nu = 3, \mu = 3, \sigma = 10) \quad (28)$$

$$\theta \sim \text{Gamma}(\alpha = \frac{1}{100}, \beta = \frac{1}{100}) \quad (29)$$

$$\pi \sim \text{Beta}(\alpha = 1, \beta = 1) \quad (30)$$

モデル6はゼロ過剰負の二項分布である。ゼロ過剰ポアソン分布の母数は4つである。最後に、検証モデルを定式化する。

検証モデル :

$$p_1(k) = \pi_2 p_2(k) + \pi_3 p_2(k) \quad (31)$$

$$p_2(k) = \begin{cases} \pi_1 & \text{if } k = 0 \\ (1 - \pi) \frac{\text{NegativeBinomial}(k|\mu_1, \theta_1)}{1 - \text{NegativeBinomialCDF}(k|\mu_1, \theta_1)} & \text{if } k > 0 \end{cases} \quad (32)$$

$$p_3(k) = \text{NegativeBinomial}(k|\mu_2, \theta_2) \quad (33)$$

$$\mu_1, \mu_2 \sim \text{Student_t}(\nu = 3, \mu = 3, \sigma = 10) \quad (34)$$

$$\theta_1, \theta_2 \sim \text{Gamma}(\alpha = \frac{1}{100}, \beta = \frac{1}{100}) \quad (35)$$

$$\pi_1 \sim \text{Uniform}(a = 0, b = 1) \quad (36)$$

$$\pi_2, \pi_3 \sim \text{Logistic}(\mu = 0, \sigma = 1) \quad (37)$$

$$\pi_4 = \pi_1 \pi_2 \quad (38)$$

$$\pi_5 = \pi_4 - \pi_1 \pi_2 \quad (39)$$

この検証モデルにおいて、(31)式は、ハードル負の二項分布と負の二項分布の混合分布であることを表している。(32)式は、ハードル負の二項分布であり、(33)式は負の二項分布である。(34)式から(37)式までは7つの母数の事前分布を表している。 π_4 および π_5 は生成量である。前者は、実際に0の値を取る確率であるからハードル1を超えないユーザの比率を、後者はハードル負の二項分布においてハードル1を超えたユーザの比率を示す。

最後に、2つ目のステップとして、コンバージョンのダミー変数(CV)がログイン回数のロジスティック関数として近似できることを示すモデルは以下のとおりである。

$$\log\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_1 k_i \quad (40)$$

$$CV_i \sim \text{Bernoulli}(p_i) \quad (41)$$

$$\beta_0 \sim \text{Student_t}(\nu = 3, \mu = 3, \sigma = 10) \quad (42)$$

$$\beta_1 \sim \text{Uniform}(\alpha = -\infty, \infty) \quad (43)$$

なお、MCMCの推定は、すべてのモデルにおいて、チェーン数を4、MCMCサンプル数を2,000、焼却区間(burn-in interval)を1,000、間引き区間なしとした。収束の判断には、モデルに含まれるすべての母数における \hat{R} が1.00であることを基準とした。

4. 結果

全モデルにおける全母数についての \hat{R} を検討したところ、すべての値が1.00の値を示したため、MCMC サンプルングは全モデルについて収束したとの判断を与えた。その後、WAICを使用して、モデルの適合度を比較した。全モデルのWAICおよびその標準誤差を表3にまとめる。次に、モデル1~6の事後期待値による予測分布を図6に示す。図7は検証モデルの予測分布である。

表3 全モデルの適合度比較

	モデルの概略	推定母数の数	WAIC	SE
モデル 1	ポアソン分布	1	13578.55	414.97
モデル 2	負の二項分布	2	6899.97	44.27
モデル 3	ハードルポアソン分布	2	12445.86	577.17
モデル 4	ハードル負の二項分布	3	6957.84	51.56
モデル 5	ゼロ過剰ポアソン分布	2	12446.36	577.24
モデル 6	ゼロ過剰負の二項分布	3	6957.96	51.55
検証モデル	ハードル負の二項+負の二項混合モデル	7	6909.91	52.73

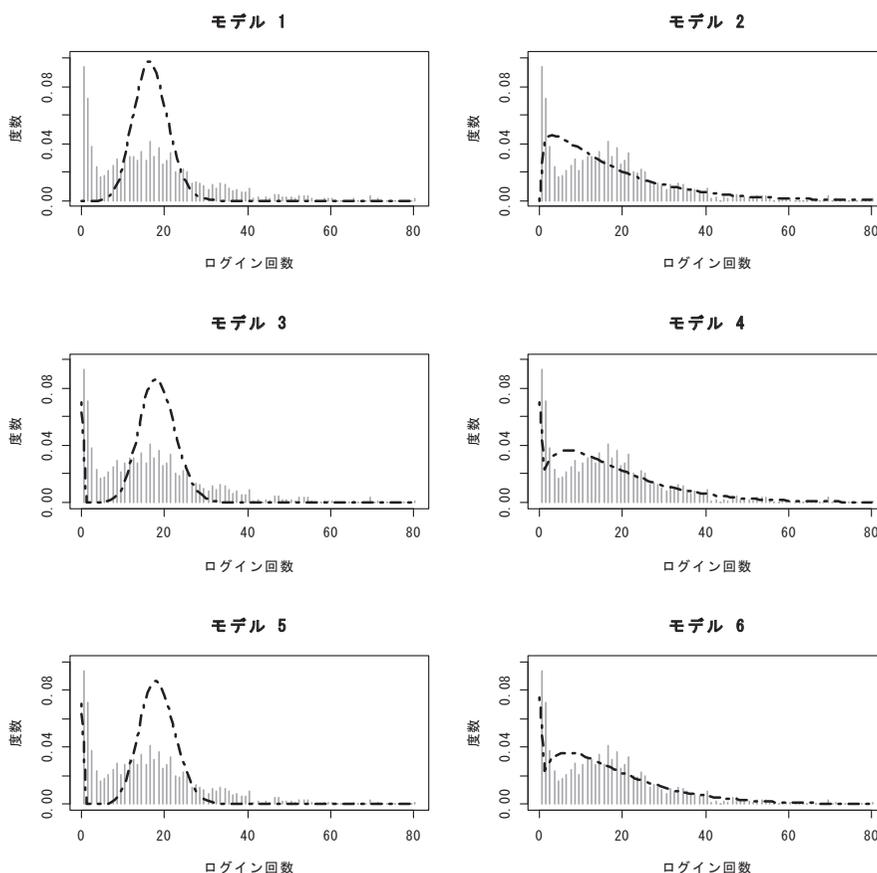


図6 観測の度数分布とモデル1~6の推定結果に基づく事後予測分布の比較⁴⁾

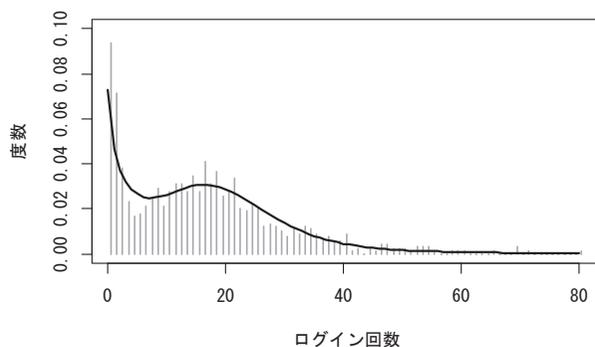


図7 観測の度数分布と検証モデルの推定結果に基づく事後予測分布の比較

WAICにおいて、最も優れた適合度を示したのは、モデル2の負の二項分布であったが、僅差で検証モデルが次に優れた適合度を示した。しかしながら、モデル2における予測分布を視認すると、確率質量関数と実測の度数が乖離していることが分かる。モデル2は母数が少なく、モデルとして儉約的であることにも注意が必要である。一方、検証モデルは、多くの母数を持つ複雑なモデルであるものの、モデル2を除けば、他よりも優れた適合度を示した上、予測分布は観測の度数分布と非常に近似している。このことは、図8の分位点-分位点プロットによって確認することができる。左側の検証モデルの方が、実測値との乖離が少ないことが視察される。

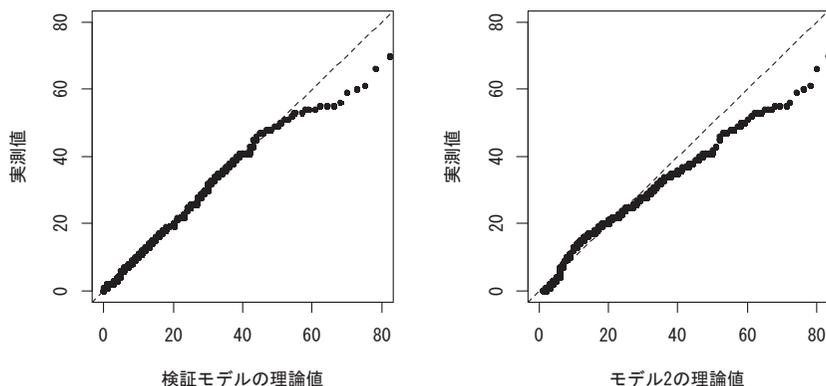


図8 観測値とモデル2および検証モデルにおける理論値の分位点-分位点プロット

一方、ハードルモデルやゼロ過剰モデルは、いずれも大差ない適合度を示しており、概して適合度が検証モデルよりも低い。特に、ポアソン分布、ハードルポアソン分布、そしてゼロ過剰ポアソン分布の適合度は低く、全体の分布形状に対してモデルが近似できていないことが示唆される。これらのことから、全体として、検証モデルは十分に優れたモデルであると評価した。

検証モデルは、ハードル負の二項分布と負の二項分布の混合分布であるため、2つの分布の概形をそれぞれ確認する必要がある。図9の実線で示される関数は、ハードル負の二項分布と負の二項分布の確率質量関数を滑らかに描いたものである。一方、破線はこのモデル全体による確率質量関数である。2つの分布はそれぞれ異なる位置と形状をもつことが確認できる。

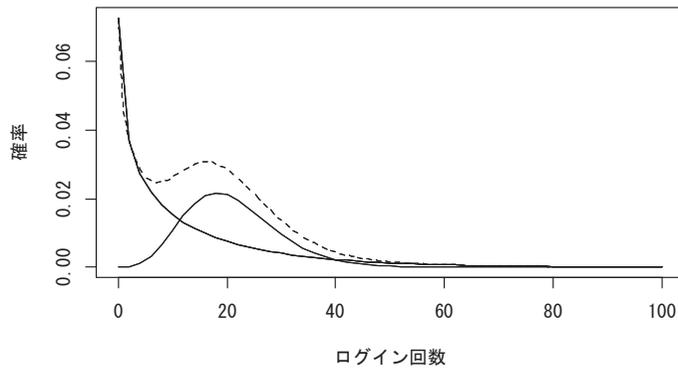


図9 混合分布の概観

次に、検証モデルにおける母数の推定結果を検討した。表4はそれぞれの母数における推定結果をまとめたものである。この母数化の下で、 e^{μ_1} および e^{μ_2} は、それぞれの分布の期待値に相当する。なおここでの e はネイピア数である。よって、ハードル2を超えないユーザのログイン回数の期待値は13.46回であり、その95%信用区間は、[9.11, 16.95]である。一方、ハードル2を超えるユーザのログイン回数の期待値は21.33回であり、その95%信用区間は、[19.69, 22.65]であった。なお、この信用区間はパーセンタイル法によって構築したものである。

図10は、 μ_1 , μ_2 , θ_1 , θ_2 における事後分布に近似するMCMCサンプルのカーネル密度曲線を示したものである。

また、 π_4 および π_5 を生成量として計算したところ、表5の結果となった。 π_4 はハードル1を超えないユーザの比率、そして π_5 はハードル2を超えないユーザの比率、そして π_3 はハードル2を超えたユーザの比率に対応する。ハードル1を超えないユーザの比率は、事後期待値において7%程度であった。ハードル2を超えたユーザの比率は、48%程度であり、残りの44%程度がハードル2を超えた計算になる。これらの推定値はそれぞれのマイクロコンバージョンにおけるCVRとして捉えることもできる。図11は、 π_4 , π_5 そして π_3 の事後分布の概観である。

表4 検証モデルにおける母数の推定結果

母数	リンク関数	事後期待値	標準偏差	95%信用区間下限	95%信用区間上限	\hat{R}
μ_1	対数	2.60	0.16	2.21	2.83	1.00
μ_2	対数	3.05	0.04	2.98	3.12	1.00
θ_0	恒等	0.74	0.18	0.39	1.09	1.00
θ_1	恒等	8.56	3.86	4.35	18.12	1.00
π_1	恒等	0.13	0.03	0.09	0.19	1.00
π_2	恒等	0.56	0.08	0.39	0.71	1.00
π_3	恒等	0.44	0.08	2.29	0.61	1.00

表5 検証モデルにおける生成量

母数	事後期待値	標準偏差	95%信用区間下限	95%信用区間上限	\hat{R}
π_4	0.07	0.08	0.06	0.09	1.00
π_5	0.48	0.08	0.32	0.64	1.00

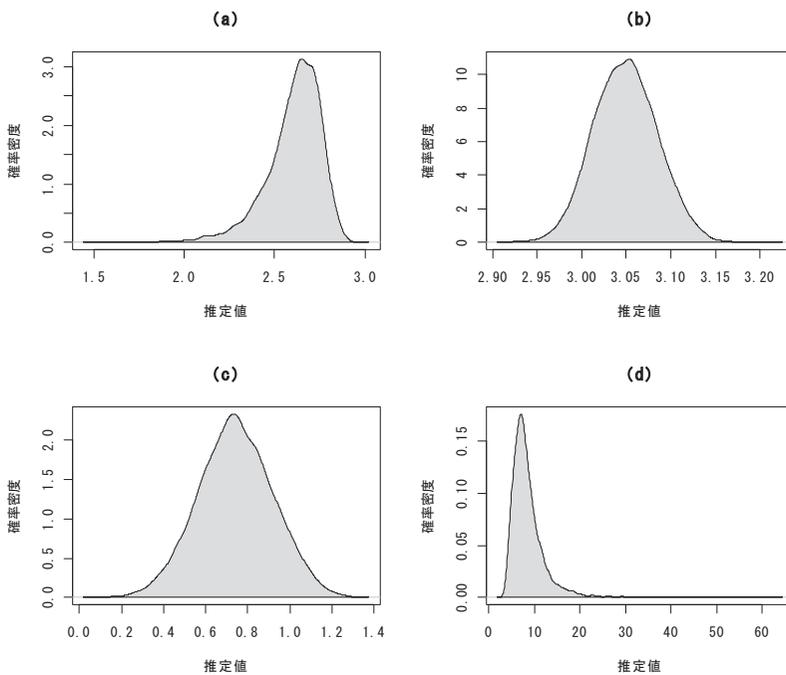


図10 各母数の事後分布の概観

注：(a)～(d) はそれぞれ μ_1 , μ_2 , θ_1 , θ_2 に対応している。

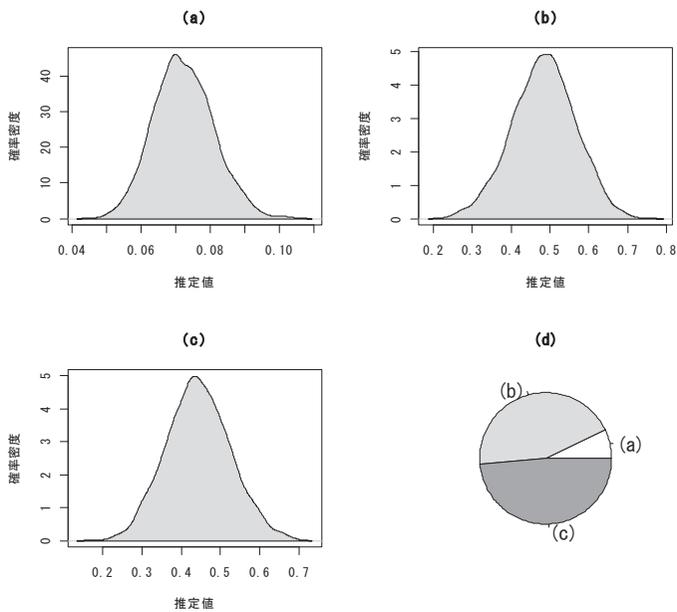


図11 π_4 , π_5 および π_3 の事後分布の概観

注：(a)～(c) はそれぞれの π_4 , π_5 そして π_3 に対応している。
 (d) は事後期待値に基づく比率を表すグラフである。

2段階目のステップとして、コンバージョンはログイン回数のロジスティック回帰モデルとして近似されるということを検証した。マルコフ連鎖モンテカルロ法によるロジスティック回帰モデルの推定結果は、表6である。図12は、切片の母数である β_0 およびログイン回数の係数である β_1 の事後分布の概観をカーネル密度曲線によって表している。

表6 ロジスティック回帰モデルにおける推定結果

母数	リンク関数	事後期待値	標準偏差	95%信用区間下限	95%信用区間上限	\hat{R}
β_0	ロジット	-2.93	0.20	-3.32	-2.55	1.00
β_1	ロジット	0.16	0.01	0.14	0.18	1.00

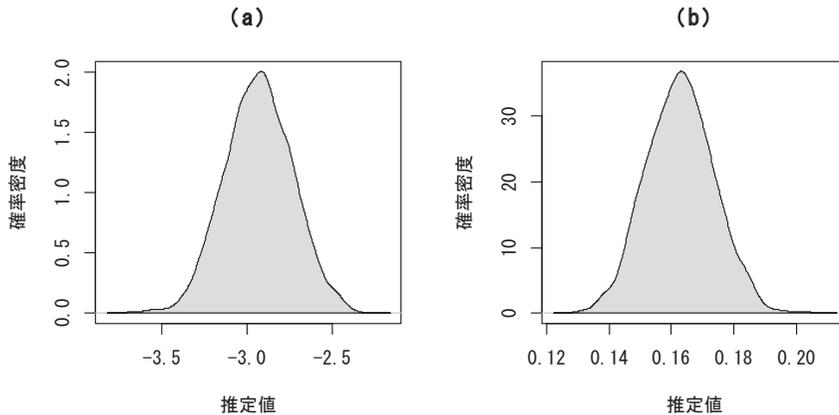


図12 β_0 (a) および β_1 (b) における事後分布の概観

β_1 の信用区間が $[0.14, 0.18]$ であり、明らかに0を跨がないことから考えても、コンバージョン確率はログイン回数が多くなるにつれ上がっていく関数の関係にあるといえる。図13は、ログイン回数とコンバージョン確率の関係を示したグラフである。このロジスティック回帰モデルにおいて、ログイン回数がおよそ18回のときのコンバージョン確率が.50に対応する。さらに、ハードル1を超えないユーザのログイン回数の期待値はおよそ13回であったが、このログイン回数に対応するコンバージョン確率は、

$$p_{cv} = \frac{\exp(-2.93 + 0.16k)}{1 + \exp(-2.93 + 0.16k)} \quad (44)$$

によって求まることから、およそ30%である。一方、ハードル1を超えたユーザの期待値はおよそ21回であったため、これに対応する確率は61%であり、2倍以上の確率をもって、コンバージョンが期待できることがわかる。

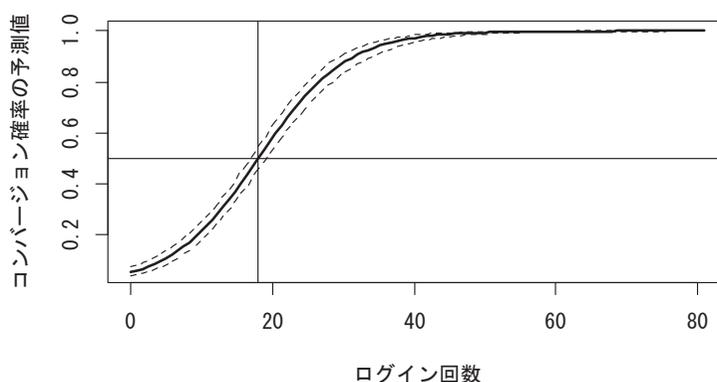


図13 モデル上におけるログイン回数とコンバージョン率

このロジスティック回帰モデルをコンバージョン予測の判定機として使用すると、単純に18回のログインの有無によって判定することになる。この予測法による混同行列が表7である。混同行列において、感度は.83、特異度は.23、陽性適中率は.81、陰性的中率は.87、そしてオッズ比は14.06であった。コンバージョン1の期待値を大きく超える18回のログインは、コンテンツの全消化であるコンバージョンを予測するための1つの目安になるかもしれない。

表7 ロジスティック回帰モデルにおける予測の混同行列

	コンバージョンあり	コンバージョンなし	計
18回以上	421	98	519
18回以下	85	295	380
計	506	393	899

5. 結論

本稿では、オンライン英語学習教材における学習の取り組みを一種のコンバージョンとして捉える見方を紹介し、さらにコンバージョンの過程であるマイクロコンバージョンという概念を、学習行動の機序として想定し、マイクロコンバージョン間の遷移をハードルとして捉える多重ハードルモデルを提案した。多重ハードルモデルの性能について検討するために、ハードル負の二項分布と負の二項分布による混合分布によってこれを実装し、この実装が比較的優れた近似を与えることをベイズ推定によって示した。さらに、この多重ハードルモデルによって近似されるログイン回数は、最終的なコンバージョンを予測する変数として有益であることを、ロジスティック回帰モデルの近似によって示した。

今後の展望として優先すべきことは、このモデルを踏まえた上でのコンバージョン最適化に関する戦略の企画立案である。現在、多くのオンライン教材の活用に関して、最終的なコンバージョンに至るマイクロコンバージョンを想定することなく、さらに、ユーザのオンライン上の行動データを十分に活用することがないまま教育実践を行っている。本稿が提案した、オンライン教材の管理運営に関してコンバージョン概念の活用や、ログイン回数などによってマイクロコンバージョン毎にユーザ比率を見積もる手法は、1つの示唆を与えるものだと考えられる。

引用文献

- Buerkner, P. C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80, 1–28.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76. DOI: 10.18637/jss.v076.i01
- Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*, 39, 829–844.
- Deb, P. & Trivedi, P. K. (1997). Demand for medical care by the elderly: A finite mixture approach. *Journal of Applied Econometrics*, 12, 313–336.
- Deb, P., & Trivedi, P. K. (2002). The structure of demand for health care: latent class versus two-part models. *Journal of Health Economics*, 21, 601–625.
- Grogger, J. T., & Carson, R. T. (1991). Models for truncated counts. *Journal of Applied Econometrics*, 6, 225–238.
- Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, 90, 36–53.
- Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34, 1–14.
- Li, L. Y., & Tsai, C. C. (2017). Accessing online learning material: Quantitative behavior patterns and their effects on motivation and learning performance. *Computers & Education*, 114, 286–297.
- 増原宏明 (2004). Finite Mixture モデルを用いた受診行動の比較分析. 『医療と社会』14. 35–50.
- McDowell, W. C., Wilson, R. C., & Kile Jr, C. O. (2016). An examination of retail website design and conversion rate. *Journal of Business Research*, 69, 4837–4842.
- McLachlan, G. J. & Peel, D. (2000). *Finite Mixture Models*. Wiley.
- Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, 33, 341–365.
- R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Stan Development Team. (2018). *Stan Modeling Language Users Guide and Reference Manual*, Version 2.18.0. <http://mc-stan.org>

注

- 1) Deb and Trivedi (1997, 2002), 増原 (2004) は, 医療需要分析において, 医療機関への受診行動について本稿と同様のモデリングを行っている。
- 2) Cragg (1971) が提案したダブルハードルモデルを示すものではない。
- 3) 本稿のモデルの基礎的な側面に関しては, Grogger and Carson (1991), Lambert (1992), Mullahy (1986), McLachlan and Peel (2000)などを参考のこと。なお, 有限混合分布やハードルモデルやゼロ過剰モデルを応用した実証研究は無数にある。
- 4) 確率質量関数をこのように曲線で表すことは, 本来適切ではない。ただし, ここでは視認性を優先した。

ABSTRACT

A Multiple Hurdle Model for Students' Conversions in Online English Learning Materials

Kunihiro KUSANAGI

Institute for Foreign Language Research and Education

Hiroshima University

The present study provides a multiple hurdle model which computationally explains complex patterns of students' engagement levels in online learning materials. It has been empirically known that online learning log data such as login counts and the number of learning contents studied by students quite often deviate from ordinary discrete probabilistic distributions such as a Poisson distribution and a negative binomial distribution. As an underlying mechanism of this empirical fact, the present study posits that there are latent sub-processes to obtain learning outcomes, which we call micro conversions, and between them, the model also assumes hurdles that students clear and fail. This modeling framework was statistically implemented as a finite mixture distribution model that mixes a hurdle negative binomial distribution and an ordinary negative binomial distribution. Using Bayesian modeling with the Hamiltonian Monte Carlo method, the model was fitted to the real login data of 899 students in Hiroshima university and achieved a relatively good fit.