Relatório

Ana Paula Silva de Souza, Caio Massaharu Damaceno Fumiya, Marcelo Torres do Ó, Matheus Augusto Rosa Silvério

Escola de Artes, Ciências e Humanidades – Universidade de São Paulo (USP)

{anasouza,caiofumiya,marcelo.torres.o, matheus_silverio}@usp.br

Abstract. *This paper describes resolutions to proposed problems through the use of Item Response Theory. 100 pairs of numbers that are parameters of 100 questions were considered, and five students with abilities* $\theta_1 = -1;\ \theta_1 = -0,5;\ \theta_1 = 0;\ \theta_1 = 0,5\ e\ \theta_1 = 1.$

Resumo. *Este relatório descreve soluções para problemas propostos através do uso da Teoria de Resposta ao Item que foram implementadas na linguagem de programação Java. Foram considerados 100 pares de números que são parâmetros de 100 questões, e cinco alunos com habilidades* $\theta_1 = -1;\ \theta_1 = -0,5;\ \theta_1 = 0;\ \theta_1 = 0,5\ e\ \theta_1 = 1.$

1. Introdução

A Teoria da Resposta ao Item (TRI) é definida como "um conjunto de modelos matemáticos que procuram representar a probabilidade de um indivíduo dar uma certa resposta a um item como função dos parâmetros do item e da habilidade (ou habilidades) do respondente. Essa relação é sempre expressa de tal forma que quanto maior a habilidade, maior a probabilidade de acerto no item" (ANDRADE; TAVERES; VALLE, 2000, p. 7).

$$f(x \mid \theta,\ a,\ b) = \frac{e^{a(\theta - b)}}{1 + e^{a(\theta - b)}}$$

A TRI permite que indivíduos submetidos a provas diferentes possam ser comparados justamente por ser voltada aos itens, e não a prova como um todo. Neste artigo 5 alunos com as respectivas habilidades: $\theta_1 = -1;\ \theta_1 = -0,5;\ \theta_1 = 0;\ \theta_1 = 0,5\ e\ \theta_1 = 1$ serão considerados para a realização dos experimentos descritos a seguir.

2. Variável Aleatória

Uma variável aleatória é uma função que associa um número real a cada evento de um experimento. A distribuição de uma variável aleatória é o conjunto composto pelas probabilidades de a variável aleatória *X* ser igual a um valor *x*, tal que *x* é um elemento da imagem de *X*.

Uma variável aleatória pode ter distribuição discreta, se seu contradomínio for o conjunto dos números inteiros, ou contínua, caso o contradomínio seja o conjunto dos números reais. Além disso, algumas distribuições são especiais, uma delas é a Distribuição Normal Padrão e outra é a Distribuição de Bernoulli.

2.1. Distribuição de Bernoulli

Uma variável aleatória de distribuição discreta é dita de Bernoulli se sua imagem for igual ao conjunto {0, 1}, isto é, se assumir apenas dois valores: 0 ou 1.

Se *p* denota a probabilidade de uma variável aleatória *X* com distribuição de Bernoulli assumir o valor 1, então é dito que *X* é uma variável aleatória de Bernoulli com parâmetro *p*.

2.2. Distribuição Normal Padrão

Uma variável aleatória contínua tem distribuição normal com parâmetros μ e σ^2, onde μ é o valor esperado (média) de uma variável aleatória X e σ^2 é a variância de X, se sua função densidade de probabilidade for dada por:

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Quando os valores de μ e σ^2 possuem, respectivamente, valores 0 e 1, a distribuição normal pode ser denominada Distribuição Normal Padrão.

Com o Teorema do Limite central, é permitido dizer que a soma de um conjunto de variáveis aleatórias independentes e igualmente distribuídas converge para uma distribuição normal. Assim, se as variáveis aleatórias $X_1, \ldots, X_n$ formam uma amostra aleatória de tamanho n de uma distribuição com média μ e desvio padrão σ ($0 < \sigma < \infty$), então para cada x

$$\lim_{n \to \infty} Pr\left(\frac{\overline{X}_n - \mu}{\sigma / \sqrt{n}} \leq x\right) = \Phi(x),$$

$$\lim_{n \to \infty} Pr\left(\frac{\sum_{i=1}^{n} X_n - n\mu}{\sqrt{n}\sigma} \leq x\right) = \Phi(x),$$

onde $\Phi(x)$ denota a c.d.f. da distribuição normal padrão.

Temos ainda por outro teorema que, sendo X uma distribuição normal com média μ e desvio padrão σ, e sendo F a c.d.f. de X, então $Z = \frac{X - \mu}{\sigma}$ tem a distribuição normal padrão para todo x e todo $0 < p < 1$.

$$F(x) = \Phi(\frac{x - \mu}{\sigma})$$

$$F^{-1}(p) = \mu + \sigma\Phi^{-1}(p)$$

3. Estimação de habilidade

A habilidade θ de um aluno pode ser estimada dada às respostas de um aluno a um conjunto de *n* questões, onde a_i e b_i são parâmetros da questão $1 \leq i \leq n$ através do método da máxima verossimilhança.

O procedimento consiste em calcular o produtório das *pdfs* do conjunto de questões e derivar o *log* do produtório a fim de igualá-lo a zero e encontrar um ponto de máximo.

$$k(x) = log(\prod_{i=1}^{n} f(x \mid \theta)) \frac{d}{d\theta} = \sum_{i=1}^{n} f(x \mid \theta) \frac{d}{d\theta} =$$

$$\sum_{i=1}^{n} [A_i \cdot [a_i - \frac{a_i \cdot e^{a_i(\theta - b_i)}}{1+e^{a_i(\theta - b_i)}}] \cdot (1 - A_i) \cdot (\frac{-a_i \cdot e^{a_i(\theta - b_i)}}{1+e^{a_i(\theta - b_i)}})]$$

O problema consiste em isolar a variável θ, fazendo necessário o uso de um algoritmo para aproximar a raiz de uma equação. O algoritmo usado foi o algoritmo de Newton-Raphson que dado executa a seguinte fórmula iterativa enquanto $|\theta_{k+1} - \theta_k| < \varepsilon$, onde ε é a precisão desejada.

$$\theta_{k+1} = \theta_k + \frac{g(x)}{g'(x)}$$

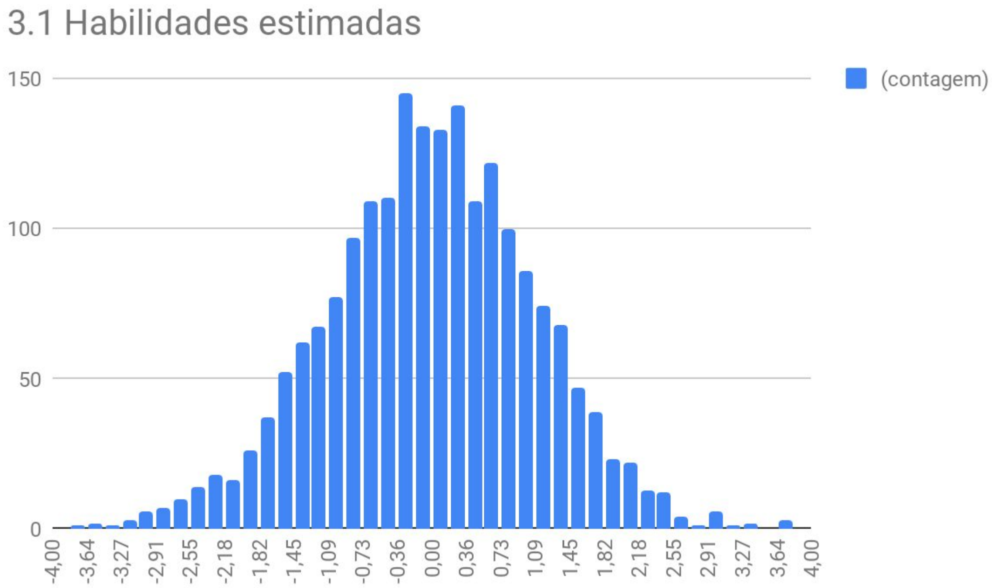
No caso foi definido $g(x) = k(x)$ e $g'(x) = k'(x)$ e o valor inicial escolhido foi $k = 0$ pois esse método exige que tal valor esteja próximo da raiz verdadeira, e uma vez estimadores foram gerador segundo uma distribuição normal padrão o valor zero foi considerado o melhor por estar dividindo uma distribuição simétrica ao meio.

Para realizar este procedimento é necessário ter as respostas do alunos às questões, já que a função $k(x)$ varia de acordo com o valor binário da variável aleatória que modela a resposta.

Em casos extremos em que o aluno acerta ou erra todas as questões é necessário atribuir a habilidade ∞ ou $-\infty$, respectivamente; pois nesses casos o valor retornado pelo método de Newton-Raphson implementando em Java será NaN (Not-a-Number), o que compromete os cálculos.

O procedimento descrito acima foi utilizado para estimar a habilidade de 2000 alunos a partir de 100 respostas geradas com uma habilidade proveniente de uma distribuição normal padrão.

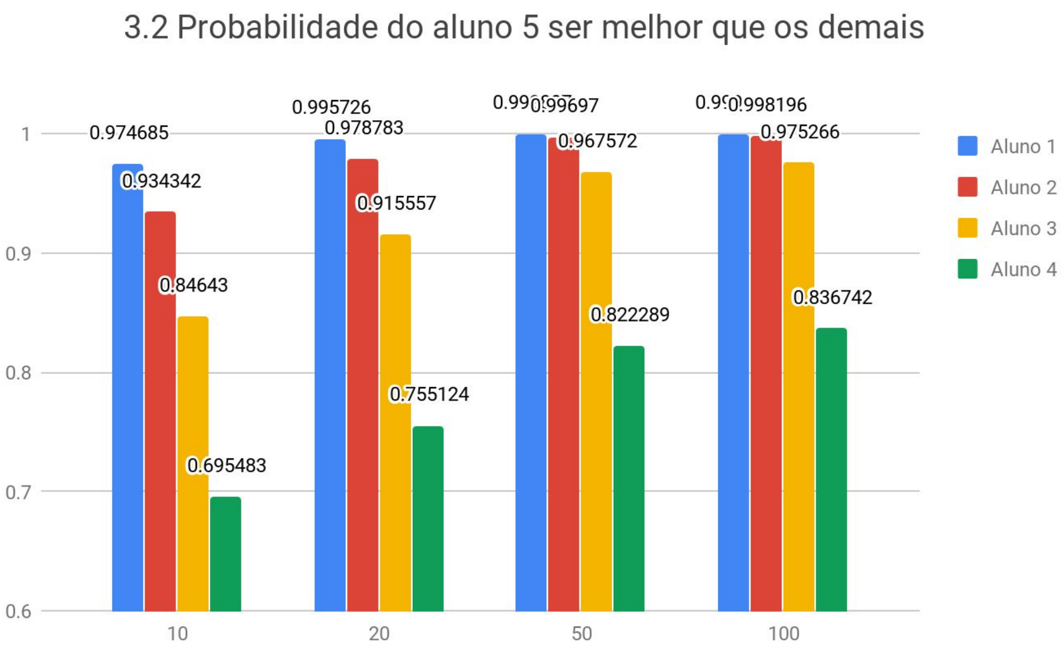
O histograma abaixo representa os valores obtidos, que se aproximam de uma normal padrão, o que evidencia a eficácia do método.



4. Seleção do melhor aluno

É possível calcular a probabilidade um determinado aluno $k \in \{1, \ldots, m\}$ ser melhor do que o outro aluno $j \in \{1, \ldots, m\} - \{k\}$ usando a habilidade estimada do aluno como nota.

A ideia consiste em simular respostas do aluno a partir de sua habilidade já conhecida e tomar como base essas respostas a questões de parâmetros já conhecidos para estimar a habilidade θ do aluno e usá-la como critério de comparação entre dois alunos.



As provas foram montadas usando as 100 questões definidas anteriormente, porém com o critério de escolher as melhores questões para o aluno 5, que tem a maior habilidade.

A seleção consistiu em ordenar as questões de forma decrescente segundo o valor de $h(\theta_5) - h(\theta_4)$, isto é, usando o critério de maior diferença de probabilidade de acerto entre o aluno 5 e o aluno 4; bastando assim selecionar as 10 primeiras questões para a prova de 10 questões e assim por diante.

Tal critério garante que a prova que favorece o aluno 5 em relação a todos os outros alunos uma vez que o aluno 4 é o aluno com a segunda maior habilidade dentre os 5 alunos em questão.

A respostas seguem uma distribuição Bernoulli com parâmetro p dado pela TRI, e por isso foram simuladas com o auxílio de uma função *rand()* que retorna um número pseudo-aleatório no intervalo $[0, 1)$. Quando $rand() < p$, é considerado que o aluno acertou a questão, caso contrário é considerado que ele errou a questão.

Analisando o gráficos dos resultados obtidos após 1.000.000 de simulações é notório que a probabilidade de o aluno 5 ser melhor do os outros alunos é proporcional ao número de questões da prova.

5. Intervalo de confiança da nota dada pela habilidade

Um intervalo de confiança com coeficiente de confiança $\gamma = 0,9$ é um intervalo com 90% ou mais das notas do aluno, que no caso são dadas pela habilidade estimada anteriormente.

Para este intervalo todas as habilidades estimadas na seção 4 foram salvas em uma estrutura de dados e foram ordenadas de maneira crescente. Então, os 5% primeiros valores e os 5% últimos valores foram ignorados, restando 90% (900.000) das habilidades estimadas.

Intervalo de confiança da nota dada pela habilidade

Questões	Aluno 1	Aluno 2	Aluno 3	Aluno 4	Aluno 5
10 Questões	-infinito a 0,0763	-1,8317 a 0,5295	-0,9585 a 0,9987	-0,3651 a 1,4983	0,1312 a 2,0918
20 Questões	-1,9417 a -0,3025	-1,3599 a 0,1836	-0,6884 a 0,657	-0,1526 a 1,1425	0,3837 a 1,666
50 Questões	-1,6592 a -0,4654	-1,0794 a 0,0049	-0,5272 a 0,4901	0,0021 a 0,9948	0,5108 a 1,521
100 Questões	-1,5918 a -0,5004	-1,0319 a -0,0267	-0,4907 a 0,4611	0,0321 a 0,965	0,5397 a 1,4856

6. Intervalo de confiança da porcentagem de acertos dado por uma Distribuição Normal Padrão

Um intervalo de confiança da porcentagem de acertos do aluno com coeficiente de confiança $\gamma = 0,9$ em uma prova com n questões também pode ser calculado definindo para representar o número de acertos do aluno a uma prova uma variável aleatória

$$Y = \sum_{i=1}^{n} X_i,$$

onde cada X_i é uma variável aleatória com distribuição de Bernoulli com parâmetro p_i, representando a resposta de um aluno a questão i. Temos então que a variável aleatória Y se aproxima de uma distribuição normal.

Segundo o teorema mencionado na seção 2 temos que a média e a variância de são respectivamente

$$\mu = E(Y) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p_i \text{ e}$$

$$\sigma^2 = Var(Y) = \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n} p_i(1-p_1)$$

Uma vez que a porcentagem de questões acertadas é igual a $\frac{Y}{n}$ temos que o intervalo de confiança desejado pode ser calculado da seguinte forma:

$$Pr(A \le \tfrac{Y}{n} \le B) = Pr(n \cdot A \le Y \le n \cdot B) = Pr(n \cdot A - \mu \le Y - \mu \le n \cdot B - \mu) =$$

$$Pr(\tfrac{n \cdot A - \mu}{\sigma} \le \tfrac{Y - \mu}{\sigma} \le \tfrac{n \cdot B - \mu}{\sigma}) \ge 0,9 \iff Pr(\tfrac{Y - \mu}{\sigma} \le \tfrac{n \cdot B - \mu}{\sigma}) - Pr(\tfrac{Y - \mu}{\sigma} \le \tfrac{n \cdot A - \mu}{\sigma}) \ge 0,9$$

$$\iff \Phi(\tfrac{n \cdot B - \mu}{\sigma}) - \Phi(\tfrac{n \cdot A - \mu}{\sigma}) \ge 0,9$$

<=> $\Phi(v_2) = \Phi(\frac{n \cdot B - \mu}{\sigma}) = 0,95\ e\ \Phi(v_1) = \Phi(\frac{n \cdot B - \mu}{\sigma}) = 0,05$, pois a distribuição é simétrica. Consultando a tabela de probabilidades da distribuição normal padrão temos que $v_2 = 1,65$ e $v_1 = -1,65$, então:

$$A = \frac{\mu - v_1 \cdot \sigma}{n} \text{ e } B = \frac{\mu + v_2 \cdot \sigma}{n}$$

Intervalo de Confiança da Nota do Aluno (em porcentagem)

Provas	Aluno 1	Aluno 2	Aluno 3	Aluno 4	Aluno 5
10 Questões	0% - 40%	10% - 50%	10% - 60%	20% - 70%	40% - 80%
20 Questões	5% - 30%	10% - 40%	20% - 55%	30% - 65%	40% - 75%
50 Questões	16% - 34%	22% - 42%	33% - 52%	38% - 60%	48% - 70%
100 Questões	21% - 41%	31% - 46%	36% - 52%	42% - 58%	48% - 78%

Referências

DeGROOT, M. H.; SCHERVISH, M. J. (2012), Probability and Statistics, Addison Wesley, 4th edition.

Dalton Francisco de Andrade, Heliton Ribeiro Tavares e Raquel da Cunha Valle. Teoria da Resposta ao Item: Conceitos e Aplicações. SINAPE, 2000..