



EACH



Escola de Artes, Ciências e Humanidades
Universidade de São Paulo

Relatório EP de Estatística - Parte 2

Leonardo Soares Santos - N°USP 10284782

Marcelo Euzebio Scarton - N°USP 10367231

Vitor Ribeiro Chrisostomo - N°USP 10387623

Introdução

Em provas de múltipla escolha, com as finalidades de minimizar os 'chutes' dos candidatos e manter um mesmo nível de dificuldade entre as provas, muitos concursos (Enem e TOEFL) utilizam o TRI (Teoria de Resposta ao Item). No TRI, são estimados o nível de habilidade dos alunos, o grau de dificuldade de cada questão, seu poder de discriminação e a probabilidade de acerto ao acaso destas. Neste EP demonstraremos, como é estimado o nível de habilidade dos alunos (θ), a partir de um arquivo *respostas.txt* dado, que contém as respostas de 2000 alunos para uma prova de 100 questões. Em seguida, será calculada a chance de um aluno 5, com um certo θ , ter uma nota maior que outros 4 alunos com diferentes θ s. Feito isso, calcularemos os intervalos de confiança das notas dos alunos de 2 maneiras diferentes.

Metodologia

• Parte 3.1 - Estimador Pontual

Nesta parte a proposta era, a partir do arquivo *respostas.txt* dado, contendo a respostas de 2000 alunos para 100 questões, estimar o teta (nível de habilidade) de cada aluno. Para isso foi utilizado o método de máximo verossimilhança, explicado a seguir. Inicialmente tínhamos a seguinte fórmula:

$$L(\theta, x) = \prod_{i=1}^n (f(x_i/\theta)) = \prod_{i=1}^{100} \left(\frac{e^{a(\theta-b)}}{1+e^{a(\theta-b)}} \right)^{x_i} \cdot \left(1 - \frac{e^{a(\theta-b)}}{1+e^{a(\theta-b)}} \right)^{(1-x_i)}.$$

Foi então aplicado log, chegando no seguinte resultado:

$$\log(L(\theta, x)) = \sum_{i=1}^{100} x_i \cdot \log\left(\frac{e^{a(\theta-b)}}{1+e^{a(\theta-b)}}\right) + (1-x_i) \log\left(\frac{1}{1+e^{a(\theta-b)}}\right).$$

Feito isso, derivamos o log:

$$\frac{dL(\theta, x)}{d\theta} = \sum_{i=1}^{100} [A_i \cdot \left[a_i - \frac{a_i \cdot e^{a_i(\theta-b_i)}}{1+e^{a_i(\theta-b_i)}} \right] + (1-A_i) \cdot \left(- \frac{a_i \cdot e^{a_i(\theta-b_i)}}{1+e^{a_i(\theta-b_i)}} \right)].$$

Então, para terminar a estimativa, temos o θ que zere a equação, e para atingir isso foi utilizado o Método de Newton-Raphson, chegando então no resultado final:

$$\frac{d^2 L(\theta, x)}{d\theta^2} = \sum_{i=1}^{100} \left[- \frac{a_i^2 \cdot e^{a_i(\theta-b_i)}}{1+e^{a_i(\theta-b_i)}} \right].$$

Para fazer o cálculo desta fórmula foi criado o método **calculaTheta** na classe *Aluno.java*, este método devolve assim que chamado, o teta estimado para tal aluno, ele é chamado na classe *Principal.java* em um laço que rodará 2000 vezes, 1 para cada aluno, e colocará os resultados no arquivo *//1.txt*.

• Parte 3.2 - Selecionando o melhor aluno - Habilidade do Aluno

Nesta parte, a nossa proposta é, a partir das questões de 10, 20 e 50 anteriormente escolhidas na parte 2.1, além da prova de 100 questões, e com a ideia de que a nota é descrita como a habilidade do aluno, calcular a probabilidade de o aluno 5, ser melhor que os alunos 1,2,3 e 4. Para isso, pegamos as 100 mil provas, de cada tipo (10,20 e 50 questões) feitas pelos alunos na parte 2.1, além da prova de 100 questões.

Os alunos então fizeram cada prova 1 vez, exceto pela prova de 100 questões, que foi feita 100 mil vezes por cada um, para fazer as provas, foram utilizados os tetas

descritos anteriormente na parte 2.1. Em seguida foram calculados novos tetas para alunos para cada uma das 100 mil provas de cada tipo, e um teta para cada execução da prova de 100 questões chamando novamente o método **calculaTheta()** para os 5 alunos. Feito o cálculo do teta, comparamos a quantidade de vezes em que o 5 foi melhor que o 1, em que o 5 foi melhor que o 2, que o 3, e que o aluno 4 para cada tipo de prova, ou seja, a quantidade de vezes em que o teta encontrado para o aluno 5 foi maior que o dos outros 4 alunos. A partir disso, com uma regra de 3, encontramos a porcentagem de vezes em que o evento de o aluno 5 ser melhor que 1, 2, 3 e 4 aconteceu, sendo esta a chance de o evento acontecer. Foram então, colocados no arquivo *//2.txt* os resultados obtidos.

● **Parte 3.3 - Intervalo de confiança - Habilidade do aluno**

Nesta parte, a proposta era, a partir das provas utilizadas acima, na parte 3.2, encontrar um intervalo de confiança com $\alpha = 0.1$ da nota de cada um dos alunos, dada pela habilidade dos alunos segundo o TRI. Para isso, são inicialmente, ordenados em ordem crescente os vetores contendo os tetas de cada aluno para cada tipo de prova.

Utilizamos então o coeficiente de confiança α , dado, para saber o tamanho do intervalo que será utilizado. No caso, em que o α é 0.1, foi utilizado um intervalo de tamanho 90% da amostra original, visto que utilizamos como amostra um intervalo de 100 mil, o nosso intervalo toma um tamanho de 90 mil tetas. Em seguida percorremos os vetores, comparando, para encontrar a menor diferença entre o início do intervalo e o fim dele. Para fazer essas comparações, foi utilizado o mesmo método **intervaloDeConfiança()** anteriormente utilizado na parte 2.3 para calcular o intervalo de confiança das notas dos alunos. Posteriormente, foram colocados os resultados obtidos no arquivo *//3.txt*.

● **Parte 3.4 - Intervalo de confiança - *t*-student**

Nesta parte, a proposta era, a partir das provas utilizadas acima, na parte 3.2 e 3.3, encontrar, assim como na parte 3.3, um intervalo de confiança com $\alpha = 0.1$, porém nesta parte, o intervalo será dado pela porcentagem de questões acertadas, sendo esta quantidade de questões acertadas aproximada por uma normal utilizando *t*-student.

Para conseguir isso, pegamos as amostras de provas, e os alunos novamente fizeram as 100 mil provas de cada tipo (10, 20 e 50 questões, e a de 100 questões 100 mil vezes), para fazer a prova utilizamos os thetas dados no enunciado da parte 2.1, em seguida guardamos os resultados obtidos na forma de porcentagem de acertos. Feito isso calculamos a média e o desvio padrão dos resultados obtido para cada tipo de prova, e ao mesmo tempo, com o α dado, através da fórmula:

$$1 - \frac{\alpha}{2}$$

é encontrado o valor 0.95, e através da tabela t -student, com grau infinito, chegamos no valor 1.645.

Com esse valor fazemos o cálculo:

$$\bar{x} \pm (\sigma^2 \cdot 1.645)$$

Sendo '+' para o fim do intervalo, e '-' para o início do intervalo. Terminado este último cálculo, colocamos os resultados no arquivo *//4.txt*.

Resultados

Parte 3.1

Na parte 3.1 os thetas foram encontrados e colocados no arquivo *//1.txt* sem problemas ou resultados que chamassem atenção. Percebendo o mesmo padrão que foi percebido com os resultados das provas feitas pelos 5 alunos na primeira parte do trabalho, no qual quanto menor o theta pior o desempenho do aluno, isso foi percebido pois quanto menos questões o aluno acertava seu theta tendia a ser menor.

Parte 3.2

Já na parte 3.2 houve resultados no qual se percebeu uma anomalia, quando aluno 1, theta igual a -1, zera a prova com certa frequência e isso causa um certo distúrbio nas probabilidades pois diferentemente da parte 2.2 na qual a probabilidade de que o aluno 5 tivesse uma nota maior que o aluno 1 é maior do que o aluno 5 comparado ao 2, nesta encaramos o cenário contrário onde o aluno 1 tinha aparentemente tirado notas maiores do que o aluno 5 mais vezes. Porém o que foi constatado é o fato de que quando o aluno zera a prova o método de Newton reage de forma estranha e o valor do theta tende ao infinito. Este resultado percebido nesta parte tem consequências nos resultados da parte 3.3 e 3.4. Foi

percebido que as probabilidades com provas de 10 e 20 questões foram menores do que as que ocorreram na parte 2.2, porém com 50 e 100 questões as probabilidades foram muito parecidas.

Parte 3.3

Na parte 3.3, em alguns casos os intervalos variaram de x a infinito. Isso com uma frequência relativamente alta para as provas de 10 questões, do aluno 1, muito raramente nas provas de 20 do aluno 1, e nenhuma vez no restante, pelo menos nos nossos testes, provavelmente se tivéssemos rodado o programa ainda mais vezes, ocorreria para o aluno 2 na prova de 10 questões. Esse resultado foi encontrado nas provas em que os alunos tiraram 0 na prova, isso explica o fato de ter ocorrido apenas para o aluno 1 (menor θ) e ter uma frequência menor de acordo com a quantidade de questões, visto que quanto maior a quantidade de questões, maior a chance de o aluno não tirar 0.

Além disto pode se perceber nesta parte que o θ inicial, aquele usado para realização da prova, pertence ao intervalo de confiança de todos os alunos mostrando coerência do método.

Parte 3.4

Como dito anteriormente a anomalia que acontece com os resultados de θ do aluno 1 na prova de questões, causa distúrbios nos resultados desta parte também, pois o início do intervalo começa com valor negativo. Apesar disso os resultados obtidos nesta parte foram muito semelhantes aos obtidos na parte 2.3.

Referências

- Gonçalves, L. (2017). *O que é Teoria de Resposta ao Item (TRI)?* - ABG Consultoria Estatística. [online] ABG Consultoria Estatística. Available at: <http://www.abgconsultoria.com.br/blog/o-que-e-teoria-de-resposta-ao-item-tri/> [Accessed 2 Jul. 2018].
- Portal.mec.gov.br. (2011). *Teoria de resposta ao item avalia habilidade e minimiza o “chute” de candidatos*. [online] Available at: <http://portal.mec.gov.br/ultimas-noticias/389-ensino-medio-2092297298/17319-teoria-de-resposta-ao-item-avalia-habilidade-e-minimiza-o-chute> [Accessed 1 Jul. 2018].

Pt.wikipedia.org. (n.d.). *Método de Newton-Raphson*. [online] Available at:
https://pt.wikipedia.org/wiki/Método_de_Newton-Raphson [Accessed 4 Jul. 2018].

Pt.wikipedia.org. (n.d.). *Teoria de resposta ao item*. [online] Available at:
https://pt.wikipedia.org/wiki/Teoria_de_resposta_ao_item [Accessed 1 Jul. 2018].