

Mateus Santos Saldanha 10882873
Raul Douglas A. Do E. Santo 10799226
Renan Ernesto Silva Pinto 10875930
Vinicius Alves Matias 10783052

Relatório EP Introdução à Estatística (Pt. 2)

São Paulo

2019

Assim como na parte 1, todo o programa é estruturado de forma que a classe EPIE2 coordene o programa e as leituras e escritas em arquivos de texto, enquanto a classe Tasks2 contém um método para resolução de cada problema, além de métodos auxiliares, cada um desses métodos principais retorna a resposta para a classe EPIE2 que então grava elas em um arquivo de texto apropriado.

II1

Inicialmente foi pensado em manualmente obter uma Likelihood, derivá-la e então obter uma função que seria usada no EP. Após verificarmos que seria inviável, a ideia foi criar um método que conseguisse fazer a derivada do produtório de um TRI, no código esse método leva o nome de derivativeProdutorio. Entretanto, o EP, por conta desse método, estava demorando para rodar. Após isso, foi visto que usando LogLikelihood as derivadas poderiam ser feitas individualmente e então somadas, então foi criado o método derivative que realiza o produtório das derivadas do log da TRI de acordo com seus parâmetros. Além disso, foi implementado o método da bissecção para encontrar o ponto crítico, que no caso da TRI é o ponto de máximo. Dessa forma, todos os tetras foram estimados de acordo com o MLE.

Na bissecção foi usado um epsilon de 6 casas decimais de precisão. O teta mínimo e máximo a serem testados na bissecção foram respectivamente -10 e 10.

Os tetras obtidos estão em torno de -5 e 5, com apenas um outlier de aproximadamente 5.6. Para teste, foi feito uma iteração de -10 a 10 ao passo de 0.1 que comparava o valor máximo da função TRI, o resultado foi que os valores obtidos coincidiram na primeira casa decimal com os obtidos através da bissecção e a derivada do logLikelihood. Também para teste, foi feito um método que aplicava provas diversas vezes aos alunos que continham tais tetras e assim foram pegos aproximadamente os valores mínimos e máximos de questões acertadas por esses alunos, depois, foi verificado a quantidade de acertos de cada aluno no arquivo de respostas para verificar se estava dentro desse range.

II2

Inicialmente foi feito a ordenação das questões de acordo com as que aumentam a probabilidade do aluno 5 ser melhor que os outros, isso foi feito comparando as probabilidades de acerto de cada questão para cada aluno. Depois, foram aplicadas provas de 10, 20, 50 e 100 questões aos alunos que retornavam as respostas, com essas respostas, foi estimado o teta de cada aluno para cada prova. Esse procedimento de aplicação das provas e estimação dos tetras foi feito 10000 vezes e a cada vez que o teta do aluno 5 era maior que o dos demais era incrementado um contador referente a cada um dos 4 alunos restantes. Após isso os valores dos contadores foram divididos por 10000 resultando na probabilidade de o aluno 5 ser melhor que os demais de acordo com a habilidade segundo a TRI.

Também, foi notado que o método da bissecção, quando a função resultante não tem pontos críticos, retornava o valor inverso do esperado, por exemplo, para um aluno que errou todas as questões, onde era esperado o menor teta possível, o método retornava o maior teta possível. Sendo assim, foi implementado uma verificação de erros, invertendo esse comportamento para o esperado.

Dados II2:

```
0.9682 0.9243 0.8348 0.6826
0.9957 0.9789 0.9153 0.755
0.9998 0.9974 0.9699 0.8283
1.0 0.9989 0.9789 0.8418
```

Dados I2:

```
0.96143 0.90412 0.78986 0.61438
0.99388 0.97085 0.88811 0.70103
0.99982 0.99624 0.96103 0.79404
```

Como pode ser visto nos dados, calcular a nota dos alunos de acordo com a habilidade segundo a TRI gera uma maior discriminação entre os alunos, sendo portanto um melhor meio de avaliação. Além disso, em ambos os casos, o conforme o aumento do número de questões, melhor é a diferenciação do aluno 5 dos demais.

II3

Foi aproveitado o problema II2 para calcular a média e a variância do teta de cada aluno em cada prova, a média, fazendo um somatório dos tetas estimados nas 10000 iterações e dividindo pela quantidade de iterações, a variância, com o valor da média em mãos e todos os tetas obtidos nas 10000, foi obtida pelo somatório do quadrado dos tetas subtraídos da média, tudo isso dividido por $10000-1$. Com esses valores a distribuição foi aproximada de uma Normal padrão disponível em uma das bibliotecas do Java, e o valor do intervalo foi calculado a partir inversa da cdf da Normal Padrão utilizando como parâmetro o grau de confiança dividido por 2, isso multiplicado pela raiz quadrada da variância sobre a raiz quadrada de 10000.

Os resultados obtidos acompanham a variância, e é visto que com o aumento da quantidade de questões da prova, menor é o intervalo. Além disso, o aluno 1 e o aluno 5, que são respectivamente o aluno com menor e maior habilidade, possuem um intervalo um pouco maior que os alunos medianos, pois existem vezes que o aluno 1 erra todas as questões e também vezes que o aluno 5 acerta todas elas, nesses casos, a derivada da função logLikelihood é sempre crescente ou decrescente, não tendo um ponto crítico e portanto sendo inviável de se estimar o teta que maximiza o likelihood através desse método. Quando isso acontece, o método da bissecção retorna o maior valor possível no caso de todas questões corretas e o menor valor no caso de todas questões erradas. Logo, a variância desses alunos aumenta e normalmente seu intervalo é ligeiramente maior que o de alunos medianos.

Com exceção do aluno 5 (por conta do problema citado), os intervalo dos alunos também diminuem conforme o aumento do teta.

II4

Primeiramente é aplicado as provas de 10, 20, 50 e 100 questões para 5 alunos e esse processo é efetuado 10000 vezes para maior precisão, e depois a aplicação é armazenada e com os resultados obtidos é calculado a média e a variância. A média é calculada com o aluno em questão junto com as respectivas provas somando-as e posteriormente às dividindo pelo número de iterações(10000) com esse cálculo é feito a variância, subtraindo a respectiva questão e aluno pela média

obtida anteriormente e com a raiz quadrada, depois disso feito é dividido pelo número de iterações menos um (por ser amostral).

Com a aplicação das provas e a média e a variância obtida desenvolve-se o intervalo de confiança com as notas dos alunos nas respectivas provas, calculada com a porcentagem de acerto das mesmas. Para criar o intervalo utiliza-se a função $z.\text{inverseCumulativeProbability}$ por α dividido por 2 e depois multiplicado pela raiz da divisão da variância pela quantidade de iterações 10000. Diferente do problema II3, no problema II4 quanto maior o número de questões maior é o intervalo de confiança com os dados obtidos, também a média e a variância acompanham esse crescimento.

Há algumas diferenças em relação à comparação com a porcentagem de acertos e o TRI (teoria de resposta ao item). A variância dos dois métodos são diferentes porém seguem um mesmo comportamento, que quanto mais questões são aplicadas a variância é diminuída em seguida, a média possui tanto valores diferentes como comportamentos diferentes, enquanto no TRI quanto mais questões aplicadas o comportamento é a média aumentar, levando em conta a porcentagem de acertos não foi notado nenhum padrão para a média. Em ambos os casos, a variância e o intervalo de confiança tendem a zero conforme o aumento do número de questões.

Intervalo de confiança da nota dos alunos 1 e 2 segundo a porcentagem de questões acertadas:

```
0.18895650231196567 0.1930434976880264 0.27101398596434234 0.27566601403565727  
0.21650079900597938 0.21950920099402418 0.29857857368576507 0.3018914263142343  
0.25185147741244246 0.2538565225875576 0.3273608460400541 0.32951915396001696  
0.3247948205778742 0.3263011794221257 0.3793178670601598 0.3808901329398373 0.
```

Intervalo de confiança da nota dos alunos 1 e 2 dada pela habilidade segundo a TRI

```
-1.5832966651000882 -1.5253164943537807 -0.7749284416126256 -0.7323870500696654  
-1.1271361160005053 -1.1017944968198816 -0.5806543538707648 -0.560220478634223  
-1.0466921950774788 -1.0317270704907777 -0.5265742794982665 -0.5130289049037225  
-1.0234292898690736 -1.0098082622969116 -0.5210859852983225 -0.5087651358292828
```

De acordo com os dados é visto que apesar de a utilização da nota dada pela habilidade segundo a TRI ter um maior poder de diferenciação entre os alunos, se comparada com a utilização da nota pela porcentagem de acertos, ela possui intervalos de confiança maiores.

Execução do EP

Para executar o EP basta executar o arquivo EPIE-1.0.jar ou pela linha de comando, ir até o diretório desse arquivo e executar a seguinte linha:

```
java -jar EPIE-1.0.jar
```

Os códigos utilizados e os arquivos .txt de saída também estão todos na mesma pasta.