Aplicação da Teoria da Resposta ao Item na gestão ambiental de municípios do Brasil

**Alunos:**

Leandro Alves da Silva - nº USP: 11868023

Thiago Ferreira Miranda - nº USP: 11925711

Segundo Anjos e Andrade (2012) diversos pesquisadores em diferentes áreas de atuação têm direcionado esforço na busca da identificação de características que possam estar presentes em indíviduos, mas não podem ser medidas diretamente (sendo estas variáveis não observadas), as quais são comumente denominadas de traço latente, e assim desenvolver de índices, escalas apropriadas para seu enquadramento. Alguns exemplos de tais características são: nível de qualidade de vida, proficiência em matemática, grau de depressão, dentre outros.

Neste tocante, duas teorias são apresentadas, a Teoria Clássica dos Testes – TCT, que utiliza o escore total do teste como sua referência de medida, e a Teoria da Resposta ao Item – TRI, cujo foco principal é o item e não o teste como um todo. Na TRI a análise de itens é efetuada por meio das estimativas de seus parâmetros e do instrumento de medida como um todo. A TRI foi desenvolvida com o propósito de resolver um problema da TCT que é a dependência da medida de proficiência em relação ao teste aplicado e dos parâmetros dos itens em relação ao conjunto dos respondentes. O que não ocorre na TRI, pois seus parâmetros (de discriminação e de dificuldade do item) não dependem do grupo de respondentes, bem como há depedência da medida de proficiência de um aluno com relação aos itens apresentados a ele. Portanto, um item mede determinada habilidade, independente de quem o está respondendo, e a proficiência de um respondente não depende dos itens que estão sendo apresentados a ele.

Maiores detalhes sobre estas duas teorias podem ser encontrados em: Gulliksen (1950), Lord e Novick (1968), Lord (1980), Vianna (1987), Pasquali (2003), Andrade, Tavares e Valle (2000), Ayala (2009), Baker e Kim (2004) e Embretson e Reise (2000).

Este trabalho tem como objetivo aplicar a teoria da resposta ao item em um caso real, para isso foi utilizado a seção de meio ambiente da Pesquisa de Informações Básicas Municipais - MUNIC e proposta a criação de um indicador para a gestão ambiental municipal por meio da Teoria da Resposta ao Item - TRI.

Como mencionado na seção introdutória, a teoria clássica dos testes tem como foco o escore obtido no teste como referência de medida utilizada na comparação entre os indíviduos participantes dos grupos observados. No entanto, na teoria clássica dos testes, além do número total de acertos, podem ser utilizadas algumas medidas para se avaliar a qualidade do instrumento de medida. Algumas dessas medidas são: o coeficiente de correlação ponto-bisserial, o coeficiente de correlação bisserial e o coeficiente alfa de Cronbach.

O coeficiente de correlação ponto-bisserial () é a correlação de Pearson entre uma variável dicotômica e o escore do teste e, é definido por:

em que,

é a média dos escores dos respondentes que acertaram o item;

é a média global dos escores do teste;

é o desvio padrão do teste;

é a proporção de respondentes que acertaram o item.

O coeficiente de correlação bisserial () é uma medida de associação entre uma variável dicotomizada e uma variável contínua e, é definido por:

em que,

é a correlação ponto-bisserial;

é a proporção de respondentes que acertaram o item;

é o valor da densidade da distribuição normal padrão no ponto em que a área da curva à esquerda deste ponto é igual a .

O coeficiente alfa de Cronbach é utilizado para medir a consistência interna do instrumento de medida e, é definido por:

em que,

é o númenro de itens;

é soma das variâncias dos itens;

é a variância global dos escores dos testes.

Esse coeficiente varia de 0 a 1. Quanto mais próximo de 0 menor a consistência e quanto mais próximo de 1 maior a consistência do teste.

A Teoria da Resposta ao Item possui modelos vários modelos propostos na literatura, neste trabalho serão abordados os modelos logísticos unidimensionais da TRI. De acordo com Andrade (2000), os modelos logísticos para itens dicotômicos são os modelos de resposta ao item mais utilizados, os quais são referidos por os modelos logísticos de 1, 2 e 3 parâmetros, que consideram, respectivamente:

O modelo logístico de 3 parâmetros é definido por:

em que,

: Representa os itens propostos para avaliar o traço latente considerado;

: Representa os indivíduos que compõe uma amostra;

: É uma variável dicotômica que assume os valores 1, quando o indivíduo responde corretamente, concorda ou satisfaz as condições do item , ou 0 caso contrário;

: Representa o traço latente do respondente j;

é a probabilidade do indivíduo com habilidade dar uma resposta correta ao item e é denominada de Função de Resposta do Item – FRI;

é o parâmetro de dificuldade (ou de posição) do item , medido na mesma escala do traço latente;

é o parâmetro de discriminação (ou inclinação) do item , com valor proporcional à inclinação da Curva Característica do Item no ponto ;

é o parâmetro do item que representa a probabilidade de indivíduos com baixa habilidade responderem corretamente o item (também chamado de probabilidade de acerto casual).

: É uma constante de escala igual a 1, mas se utiliza o valor 1,7 quando se quer que os valores da função logística se aproximem da função ogiva normal.

O modelo logístico de 2 parâmetros é semelhante ao modelo de 3 parâmetros, mas não inclui o parâmetro de acerto casual no modelo.

A função de informação do item (item information function) permite analisar o quanto um item contém de informação sobre a medida de habilidade. Ela indica a quantidade de informação que um item apresenta dentro da escala de habilidade.

Para o modelo logístico de 2 parâmetros a função de informação do item é definida como:

em que,

é a informação fornecida pelo item no nível de habilidade ;

;

;

é o parâmetro de discriminação (ou inclinação) do item .

Dada a independência entre os itens, a função de informação do teste (teste information function) é a soma das informações fornecidas por cada item, que foram calibrados em uma mesma escala. A função de informação do teste é escrita como:

O processo de estimação dos parâmetros dos itens na TRI é chamado de calibração, nessa etapa é necessário que sejam satisfeitas algumas suposições, como:

* As resposta provenientes de indivíduoes diferentes são independentes;
* Os itens são respondidos de forma independente por cada indivíduos, fixada a sua proficiência, isto é, a proficiência não é alterada no momento em que o indivíduo está fazendo o teste, sendo ela a única informação individual necessária para que um indivíduo acerte ou erre o item.

O método da Máxima Verossimilhança Marginal consiste em obter os valores dos parâmetros que maximizam a probabilidade de obter-se o conjunto de dados observado. Ou seja, com grande probabilidade têm sido descobertos os parâmetros que geraram a base de dados.

O processo é construído em duas etapas. Primeiro supõe-se alguma distribuição para as habilidades, como normal , por exemplo normal com e . Tem-se a função de verossimilhança e sua log-verossimilhança:

A partir de então a equação de verossimilhança fica apenas em função dos parâmetros dos itens e segundo o desenvolvimento de Baker (1992), Andrade, Tavares & Valle (2000) e Azevedo (2003), chega-se as seguintes equações de estimação:

onde,

é a função densidade condicional que tradicionalmente fornece estimativa da habilidade, dado o vetor de respostas do indivíduo .

As Equações 10, 11 e 12, por não terem soluções explicitas, são maximizadas através de algum método iterativo, tais como Algoritmo EM (Esperança-Maximização), “Scoring” de Fisher ou Newton-Raphson para obtenção dos valores (, , ), para cada item, que maximizam a verossimilhança gerada pelos dados (ANDRADE, TAVARES e VALLE, 2000).

A função apresentada em (13) é extremamente importante, pois ela representa a adaptação da função densidade caso seja obtido o conjunto de resposta do indivíduo j, representados por: . Em suma, cada indivíduo tem sua própria função densidade, que será usada para estimação da proficiência.

A estimação de pela média da posteriori (ou Esperança a posteriori) de , consiste em obter a esperança dessa distribuição, que pode ser escrita como:

A Pesquisa de Informações Básicas Municipais é realizada pelo Instituto Brasileiro de Geografia e Estatística - IBGE desde 1999 e, conforme publicado em seu livro, trata-se de pesquisa institucional e de registros administrativos da gestão pública municipal. O principal interesse da Munic é a gestão pública na esfera de administração, no que se refere à organização das prefeituras, bem como o quadro funcional, estrutura e políticas públicas setoriais no âmbito das áreas pesquisadas. Em 2020, a pesquisa investigou informações sobre recursos humanos, habitação, transporte, agropecuária, meio ambiente, gestão de riscos e de desastres, e COVID-19. O questionário de meio ambiente é dividido nas seguintes seções:

1. Órgão gestor do meio ambiente;
2. Recursos humanos na área do meio ambiente por vínculo e capacitação;
3. Conselho e Fundo municipal de Meio Ambiente;
4. Legislação ou instrumento de gestão ambiental;
5. Plano de Gestão Integrada de Resíduos Sólidos;
6. Programas em parceria com o Governo Federal;
7. Pagamento de Serviços Ambientais;
8. Impacto Ambiental e/ou processo/ação que resulte em impacto no ambiente.

Composto pelo seguintes itens em cada seção:

Foram adotados alguns critérios de seleção de municípios, o primeiro deles foi o de participação no questionário, algumas prefeituras obtaram por se abster da participação, o segundo foi de respostas aos itens do questionário, foram considerados os municípios com respostas “Sim” em pelo menos 5 questões. Também foram selecionados os itens que estão em acordo com os pressuposto do modelo, ou seja, foram removidos da análise itens não dicotômicos e itens que apresentavam respostas condicionais, ou seja, itens que necessitam de uma resposta prévia. Após estas remoções, restaram para a análise 4775 municípios e 60 itens.

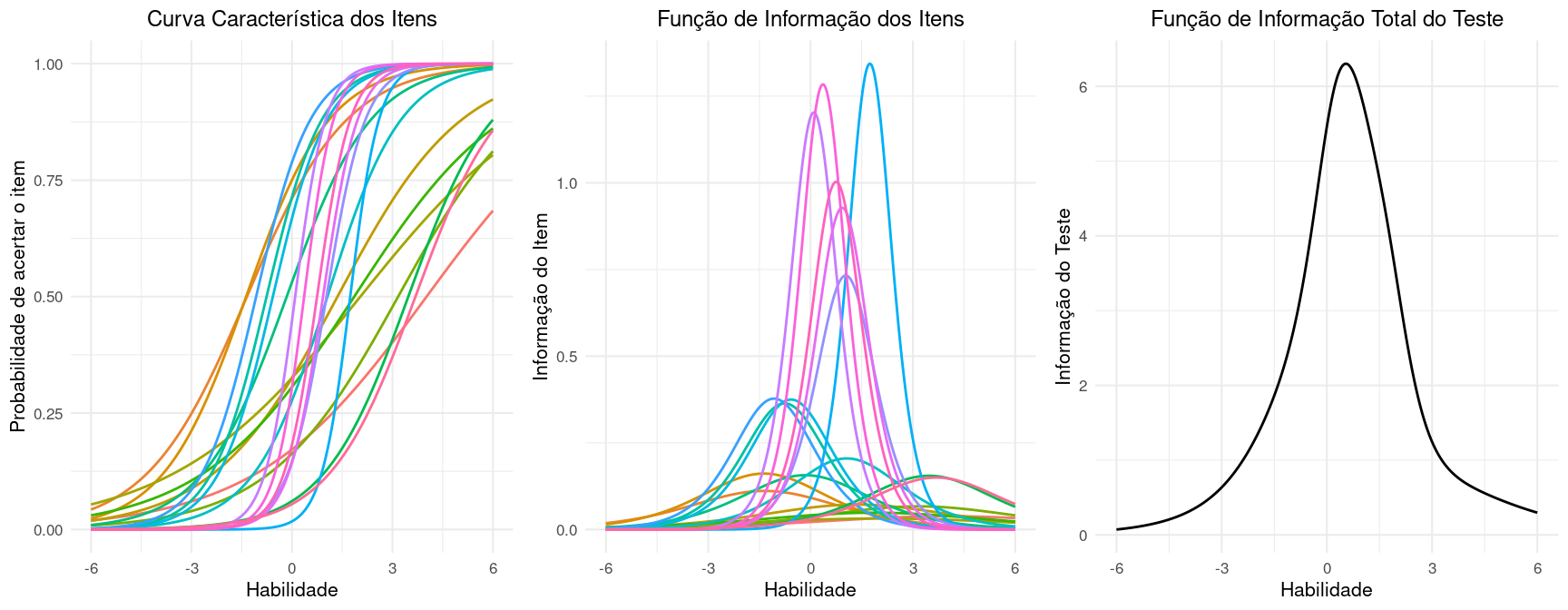
Para uma avaliação inicial dos itens e também do questionário com um todo, realizou-se a análise via teoria classica dos testes (TCT) para fazer a checagem do instrumento de avaliação, onde verificou-se o coeficiente de correlação ponto-bisserial e o alpha de cronbach. Nessa avaliação, itens podem ser removidos da análise TRI por terem coeficiente de correlação ponto-bisserial negativo, indicando uma correlação negativa entre as respostas do item em relação ao escore total do teste (TCT) e por terem alpha de cronbach alto quando excluímos o item, indicando que a exclusão dele aumentaria, de modo significativo, a consistência do teste. Portanto, para a análise pela TRI, não removeu-se nenhum item. Na tabela 9 e 10 temos as proporções de respostas e as demais estatísticas supracitadas.

A primeira etapa da TRI é a calibração dos parâmetros dos itens, foi executado o modelo logístico unidimensional de 2 parâmetros com o suporte computacional do softaware R, utilizando o pacote irtoys (Partchav e Maris,2022). Ao longo da etapa de calibração são feitas rodadas onde os itens são estimados e avaliados segundo seus respectivos parâmetros, em cada rodada são removidos da análise os itens que não se enquadram nos seguintes critérios de exclusão: i) apresentar parâmetro de discriminação e ii) apresentar parâmetro de dificuldade e/ou .

Para a calibração realizada com os dados de gestão ambiental da MUNIC, após 3 rodadas, restaram 20 itens com os seus respectivos parâmetros apresentados na tabela 11.

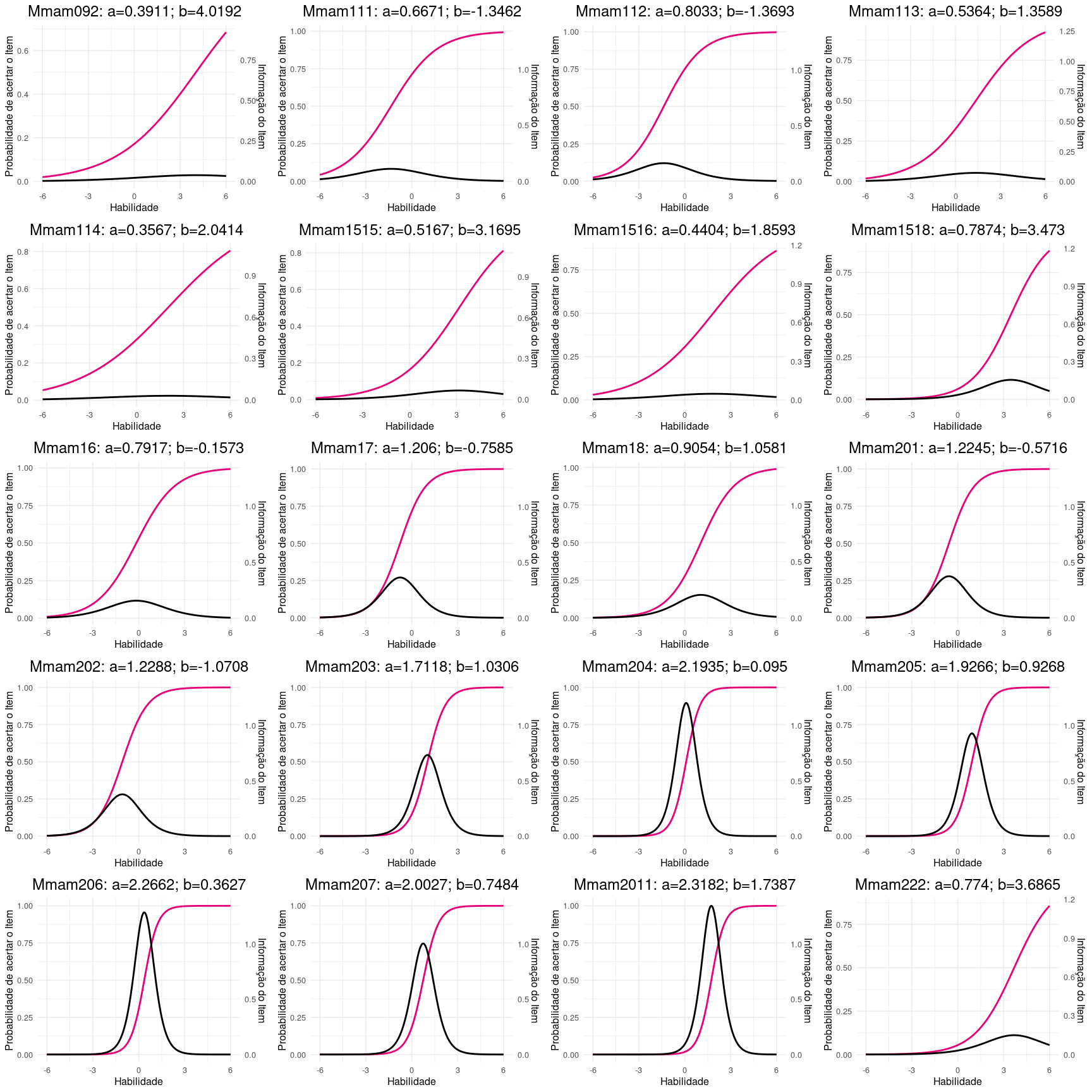
Uma vez definidos os itens e seus respectivos parâmetros, torna-se possível a elaboração e análise de outras características dos itens, tais informações são mostradas nos gráficos das Curvas Características dos Itens, de Funções de Informação dos Itens e de Função de Informação Total do Teste.

Quando o posicionamento das Curvas Características dos Itens mostradas na Figura 1 é avaliado, percebe-se que os itens não estão bem distribuídos pela escala de habilidade, pois há ausência de itens posicionados mais próximos ao início da escala (limites inferiores da escala), este evento pode ser melhor visualizado nos gráficos de Funções de Informação dos Itens, uma vez que estas curvas são centralizadas em torno dos valores de seus respectivos parâmetros de dificuldade , que por sua vez é medido na mesma escala de habilidade .



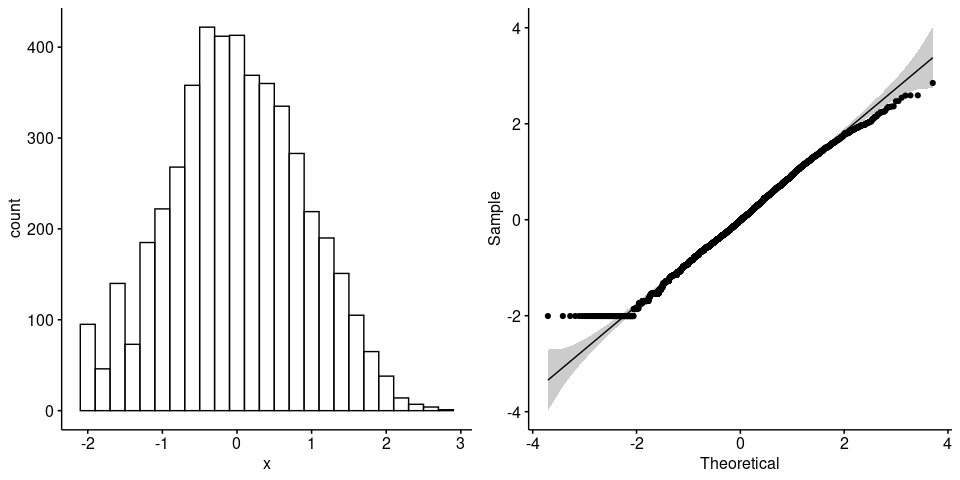
Curvas Característica e Funções de Informação dos Itens e Função de Informação Total do Teste.

Na Figura 2 temos os gráficos das Curvas Características dos Itens e de Funções de Informação dos Itens dos 20 itens que permaneceram na análise TRI apresentados um a um. Com isso é possível verificar aspectos particulares de cada item, tais como itens que possuem elevado índice de dificuldade, mas baixo potencial de discriminação como os itens Mmam092 e Mmam222. Contudo, a análise gráfica e disposição das curvas torna ainda mais evidente a ausência de itens mais próximos aos limites inferiores da escala.



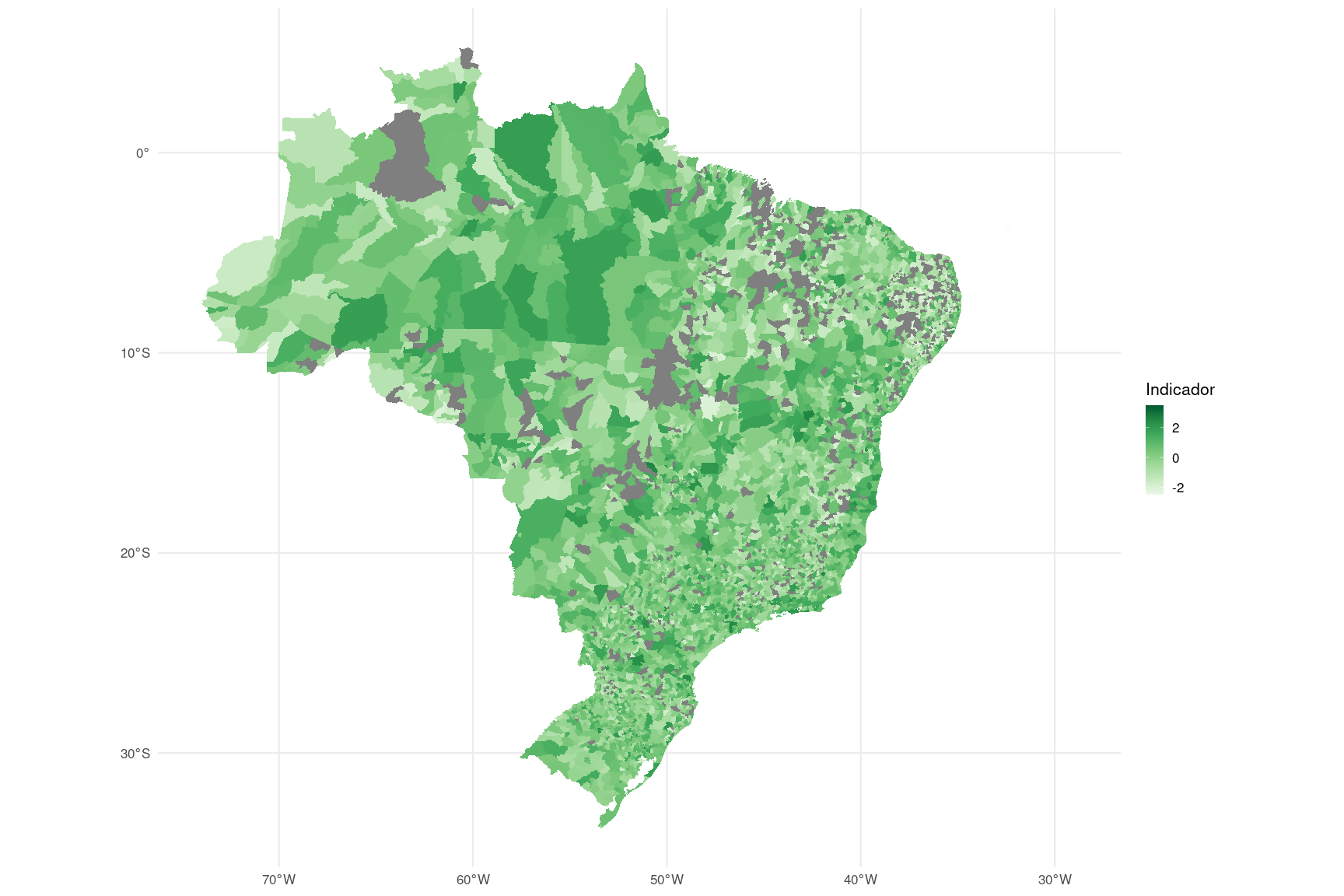
Curvas Característica e Funções de Informação dos Itens.

Em posse dos parâmetros dos itens, podemos realizar a estimação das habilidades dos municípios, esta estimação assume como premissa além da unidimensionalidade, a distribuição Normal das habilidades. A Figura 3 mostra o histograma e Q-Q Plot dos escores obtidos pela TRI, nestes constatamos que o formato da distribuição obtida é, aproximadamente, Normal, pois apresenta muitos munícipios posicionados na calda inferior esquerda da curva. Isso pode ser resultado da ausência de mais itens (questões) com baixo nível de dificuldade nesse nível da escala.



Histograma e Q-Q Plot dos escores.

No mapa abaixo temos ilustrado o indicador de gestão ambiental contruído para cada município, a partir da escala em verde podemos observar quais municípios tem um alto índice de gestão (verde escuro) e quais tem um baixo índice de gestão (verde claro). Com este indicador torna-se possível a aplicação de políticas públicas nos municípos com baixo índice de gestão ambiental, assim municípios com alto índice podem ser usados como refêrencias, dado suas respostas a itens chaves do questionário.



Mapa do indicador de gestão ambiental municipal no Brasil.

Este trabalho teve como objetivo aplicar a teoria da resposta ao item em um caso real, para isso foi utilizado o questionário de meio ambiente da Pesquisa de Informações Básicas Municipais - MUNIC e criado um indicador para a gestão ambiental municipal por meio da Teoria da Resposta ao Item - TRI. A TRI se mostrou uma boa ferramenta para a criação de indicadores, apesar do surgimento de algumas dificuldades relativas ao atendimento de alguns pressupostos, como a normalidade de escores estimados. Nota-se que uma das vantagens é relacionamento dos parâmentros dos itens e escores gerados, o que proporciona maior assertividade na orientação de políticas públicas que visem o aprimoramento da gestão ambiental nos municípios. Tendo em vista que a informação provida por cada item possibilita a formulação de iniciativas e estratégias pontuais ligadas à realidade de cada município, não estando necessariamente vinculadas ao escore geral do respondente. Mediante aos resultados obtido na estimação dos parâmetros e escores dos respondentes é percebido que se torna prematura a emissão de um parecer quanto à dimensão do instrumento (número de itens) e/ou estratificação das áreas chaves a serem incorporadas ao instrumento, visto que ainda há um vasto campo de oportunidades de análises possíveis a serem efetuados. Portanto, como trabalhos futuros podemos citar a análise TRI a partir de modelos multidimensionais, envolvendo multiplos traços latentes ou ainda modelos unidimensionais para multiplos grupos, estudando o efeito dos itens em diferentes populações.