



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA
DEPARTAMENTO DE MATEMÁTICA APLICADA



GUSTAVO DE ALMEIDA LEIVA

Estudo sobre os efeitos da pandemia do Coronavirus nos alunos do IMECC

Campinas
11/01/2021

GUSTAVO DE ALMEIDA LEIVA

**Estudo sobre os efeitos da pandemia do Coronavirus nos alunos
do IMECC**

Monografia apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos para obtenção de créditos na disciplina Projeto Supervisionado, sob a orientação do(a) Profa. Mariana Rodrigues Motta.

Resumo

Em maio de 2020, o Instituto de Matemática, Estatística e Computação Científica (IMECC) da Universidade Estadual de Campinas (UNICAMP), com o objetivo de investigar a percepção dos alunos com respeito aos desfechos da pandemia por Coronavírus e seus efeitos na saúde mental e física destes sujeitos aplicou um questionário junto aos seus alunos do instituto. Um estudo a partir de nuvem de palavras foi utilizado para investigar sobre os sentimentos reportados pelos alunos de graduação e pós-graduação. Entre os alunos que responderam ao questionário, os dados mostram que os sentimentos são similares entre os grupos, porém diferem quanto à intensidade. Posteriormente foi realizado um estudo a partir de um modelo para respostas gradativas. Esse modelo considerou todas as perguntas do formulário aplicado, e através das estimativas do ajuste foi possível acessar a importância de cada item do questionário na identificação dos fatores da pandemia que afetaram a saúde física e mental dos alunos do IMECC.

Abstract

In May 2020, the Institute of Mathematics, Statistics and Scientific Computing (IMECC) of the State University of Campinas (UNICAMP), with the objective of investigating students' perceptions of the outcomes of the Coronavirus pandemic and its effects on mental and physical health, applied a survey on the topic. A word cloud study was used to investigate the feelings reported by post-graduate and graduate students. Data shows that feelings are similar between groups, but differ in intensity. Subsequently, a study was carried out using a graded response model, with the objective of assessing the probability of the subject choosing a particular category of the applied question (item). This model considered all the questions in the applied form, and through it it was possible to identify those that discriminate the subjects.

Sumário

1	Introdução	6
2	Banco de dados	6
3	Metodologia	7
3.1	Modelando os dados através do modelo de respostas graduais	7
3.1.1	Comparando modelos	8
3.1.2	Interpretando os gráficos do modelo de respostas graduais	9
4	Resultados	10
4.1	Análise descritiva	10
4.2	Resultados da estimação do modelo de respostas graduais	13
4.2.1	Diagnóstico	14
4.2.2	Análises gráficas	15
5	Conclusão	19
6	Apêndice	21
6.1	Colunas	21

1 Introdução

Dentro de uma universidade, um assunto muito tratado e debatido é o impacto do curso de graduação e pós-graduação na saúde mental de seus alunos. Por exemplo, os fatores que afetam a qualidade do curso, bem como a qualidade da saúde mental do aluno, são aqueles relacionados à qualidade das aulas dadas, ao tratamento que os professores dão aos alunos, à pressão social para se formar no tempo mínimo do curso.

O ano de 2020 foi marcado pela pandemia do Sars-CoV-2, que levou a suspensão das aulas presenciais da UNICAMP, gerando incertezas com relação ao momento em que as aulas voltariam ao normal, como os professores lidariam com as avaliações e como estudar e aprender durante um período tão conturbado e impactante na história do mundo.

Diante deste cenário, o IMECC aplicou em maio de 2020 um questionário não obrigatório com seus alunos de graduação e pós-graduação para acessar quais impactos e fatores influenciavam na saúde mental dos respondentes. Essa monografia visa estudar as respostas dos alunos de graduação e pós-graduação do IMECC, traçando comparações entre grupos e respostas de diferentes perguntas.

Na seção Banco de dados é feita uma descrição da coleta e estrutura do banco de dados. Na seção Metodologia é descrito o modelo de respostas graduais utilizado para ajustar as respostas ao longo das perguntas do questionário. Em Resultados é apresentado os resultados das análises descritivas dos dados bem como do ajuste do modelo. Finalmente, em Conclusão descreve-se os principais achados do estudo.

2 Banco de dados

O questionário aplicado entre os alunos do IMECC não foi obrigatório. No entanto, houveram 314 respostas, 107 de alunos de pós-graduação e 207 de graduação. No anuário estatístico (da Silva et al. [2020]) da Unicamp de 2020 (base de 2019) consta 882 alunos de graduação e 541 alunos de pós-graduação. Desta forma, considerando que não houve grande variação para o ano de 2020, aproximadamente 23% dos alunos de graduação e 20% dos alunos de pós-graduação responderam ao questionário. A lista de perguntas se encontra na Tabela 1. As perguntas de interesse estão na escala Likert de 1 a 5, tal que o

valor 1 representa menor intensidade e o valor 5 maior.

Tabela 1: Codificação para perguntas

Pergunta
Já se adaptou ao isolamento social?
Enfrenta problemas financeiros?
Tem conseguido estudar satisfatoriamente?
Sente falta das atividades presenciais?
Preocupa-se com sua saúde?
Preocupa-se com sua família?
Tem se alimentado bem?
Tem dormido com qualidade?
Está sedentário?

3 Metodologia

Para estudar os dados obtidos através do questionário aplicado, contamos primeiramente com uma análise descritiva dos dados. Para investigar a influência de cada pergunta na saúde mental e física do aluno, foi aplicado o modelo de respostas graduais. Tal metodologia é descrita a seguir.

3.1 Modelando os dados através do modelo de respostas graduais

O Modelo de Respostas Graduais, ou MRG, é uma variação do modelo de Teoria de Resposta ao Item para o caso politômico, ou seja, quando a pergunta (item) tem mais do que duas categorias (Samejima [1968]). A ideia é medir um traço latente, que é uma variável não observável e mensurável associada ao sujeito j . Neste estudo, o traço latente representa o estado ou saúde mental do aluno. A partir do modelo MRG deseja-se avaliar o grau de importância dos itens (perguntas) de um formulário e qual a probabilidade do aluno escolher certa categoria de um item, dado o traço latente.

No modelo MRG, usando um modelo logito, a probabilidade $P_{i,k}^*$ de um indivíduo j escolher categorias iguais ou maiores do que k para o item i é dada por:

$$P_{i,k}^*(\theta_j) = \frac{1}{1 + e^{-a_i(\theta_j - b_{i,k})}} \quad (1)$$

onde $k = 0, \dots, m_i$ é o score ordenado, indexado pelo item i , $i = 1, \dots, I$, j indexa cada sujeito, θ_j representa parâmetro de habilidade do indivíduo j , $b_{i,k}$ representa o parâmetro de dificuldade (variável latente) da k -ésima categoria do item i e a_i é o parâmetro de discriminação (ou de inclinação) do item i .

Desta forma, a partir do modelo em (1), a probabilidade do sujeito j ter como resposta a categoria k para o item i é dada por

$$P_{i,k}(\theta_j) = \frac{1}{1 + e^{-a_i(\theta_j - b_{i,k})}} - \frac{1}{1 + e^{-a_i(\theta_j - b_{i,k+1})}}. \quad (2)$$

O modelo em (2) pode ser ajustado de forma restrita e irrestrita. Na forma restrita, a_i (parâmetro de discriminação) é fixo para todos os itens i . Na forma irrestrita, a_i é diferente para cada item i . Note que as probabilidades em (2) são calculadas a partir do valor de θ_j .

O objetivo é estimar a_i e $b_{i,k}$ (características do item) e também θ_j (característica do indivíduo - habilidades para todo i e j). O parâmetro θ_j não é observável e tratado como uma variável aleatória com distribuição normal. Seu valor é predito no processo inferencial.

3.1.1 Comparando modelos

Para escolher entre o modelo irrestrito e restrito, utilizou-se o critério de Akaike (AIC), que é um estimador do erro de predição desenvolvido em Akaike [1978] com base na teoria da informação. Ele funciona de forma comparativa para a escolha de um modelo. Quanto menor o AIC, melhor é o modelo. Além do AIC, utilizou-se o teste da razão de verossimilhança (TRV). A estatística do teste (LRT) é dada por:

$$LRT = -2(\log.Lik(\text{modelo irrestrito})) + 2(\log.Lik(\text{modelo restrito})) \quad (3)$$

onde $\log.Lik$ é o logaritmo da razão de verossimilhança de cada modelo. Quanto maior a estatística, mais probabilidade do modelo irrestrito ser o melhor.

3.1.2 Interpretando os gráficos do modelo de respostas graduais

A partir das estimativas dos parâmetros do modelo de resposta gradual é possível obter o traço latente do sujeito j (θ_j) e a probabilidade do sujeito j escolher a resposta k no item i , como descrito em (2). Feito isso, é possível construir o gráfico das curvas características para cada possível categoria do item i a partir do traço latente do sujeito j . A curva característica da categoria k do item i informa como se comporta a probabilidade (eixo y) a medida que o traço latente (eixo x) varia.

O gráfico de curva de informação indica o quanto de informação um determinado item traz com respeito à variação do traço latente.

Finalmente, o gráfico de informação do modelo indica para qual intervalo do traço latente o teste funciona melhor.

Para ser possível interpretar de forma apropriada o traço latente dos itens do questionário, fez-se uma divisão dos itens entre itens de natureza positiva e negativa, e um modelo de resposta gradual foi ajustado para cada cenário. Para os itens de natureza positiva o traço latente é interpretado como a habilidade de lidar com a pandemia e no caso dos itens de natureza negativa o traço latente é interpretado como a inabilidade de lidar com a pandemia.

Para o desenvolvimento desse estudo, foi utilizado o programa R e a IDE RStudio, com a ajuda dos pacotes de manipulação de bancos de dados e criação de gráficos *tidyverse* (Wickham et al. [2019]) e o pacote para ajustar o modelo de respostas graduais, foi utilizado o pacote *ltm* (Rizopoulos [2006]).

4 Resultados

4.1 Análise descritiva

A Figura 1 mostra a distribuição da frequência relativa da resposta (categórica) de várias das perguntas de interesse do questionário. Todas essas perguntas estão em escala de 1 a 5 (Discordo Fortemente, Discordo, Neutro, Concordo, Concordo Fortemente). Vemos que a maior parte das respostas são negativas ou neutras. Vide, por exemplo, o gráfico a respeito da preocupação com a família. A maior parte das pessoas se preocupam num nível muito alto. Como o único indicador que se mostra um pouco mais promissor é o de enfrentar problemas financeiros, nota-se que questões financeiras não é uma preocupação dos alunos, já que a tendência das respostas é positiva. Podemos observar também que a distribuição das respostas dos alunos de pós graduação e dos alunos de graduação é similar, porém entre os alunos de graduação as frequências são sempre maiores para todas as perguntas.



Figura 1: Distribuição da frequência relativa de cada categoria respondida por item.

A Figura 2 mostra como foi a experiência e o aprendizado dos alunos nas aulas remotas. Nota-se que experiência está em torno da média, com poucas pessoas nos

extremos de ambas as categorias. Vemos também que a distribuição relativa das respostas entre alunos de pós-graduação e graduação parece ser bem semelhante. No entanto as frequências das respostas dos alunos de graduação ao longo das categorias é maior do que aquelas dos alunos de pós-graduação, respectivamente.

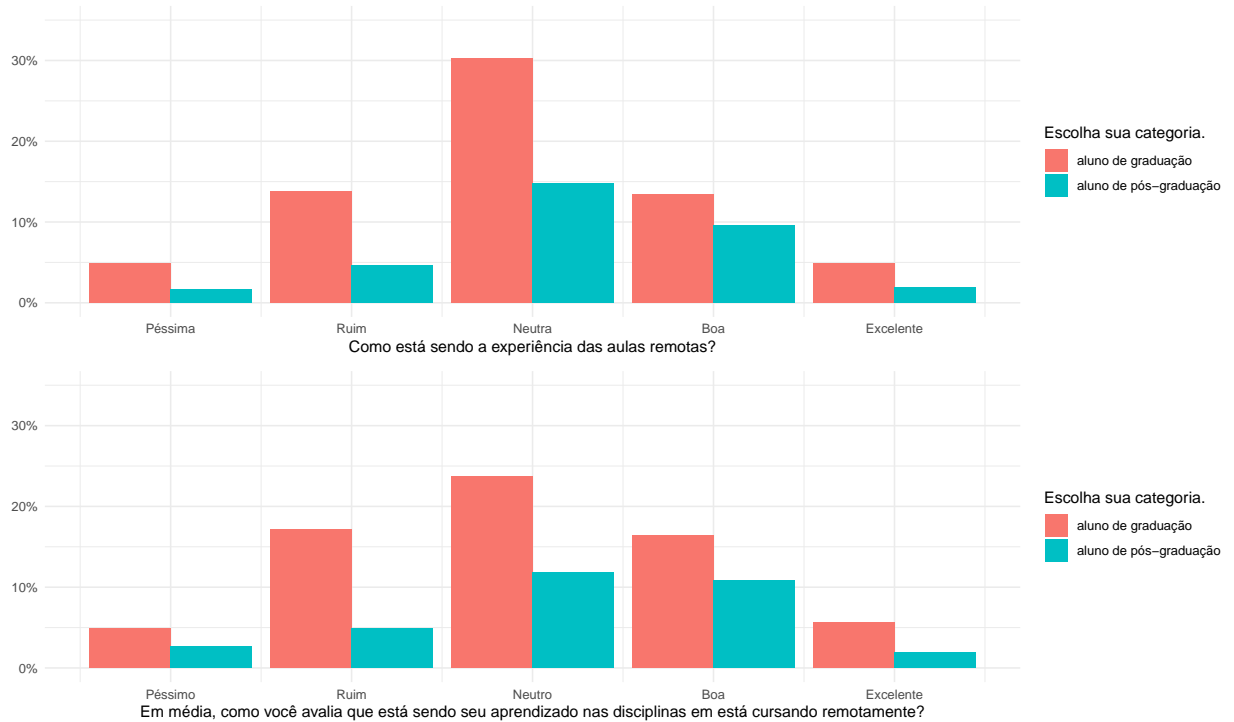


Figura 2: Distribuição da frequência relativa da experiência de aprendizado dos alunos

Uma das questões do questionário pedia que o aluno indicasse a primeira palavra que ocorresse em sua mente sobre a situação da pandemia devido a COVID-19. Uma maneira de explorar as respostas de forma visual é utilizar uma nuvem que destaca as palavras que aparecem com mais frequência, como mostra a Figura 3, segundo alunos de graduação e pós-graduação. Algumas palavras aparecem nas duas nuvens como mais comuns, como “preocupação” e “incerteza”. Porém, palavras como “medo” e “tristeza” são mais frequentes (proporcionalmente) entre os alunos da graduação, já “angústia” é mais comum entre os alunos de pós-graduação.

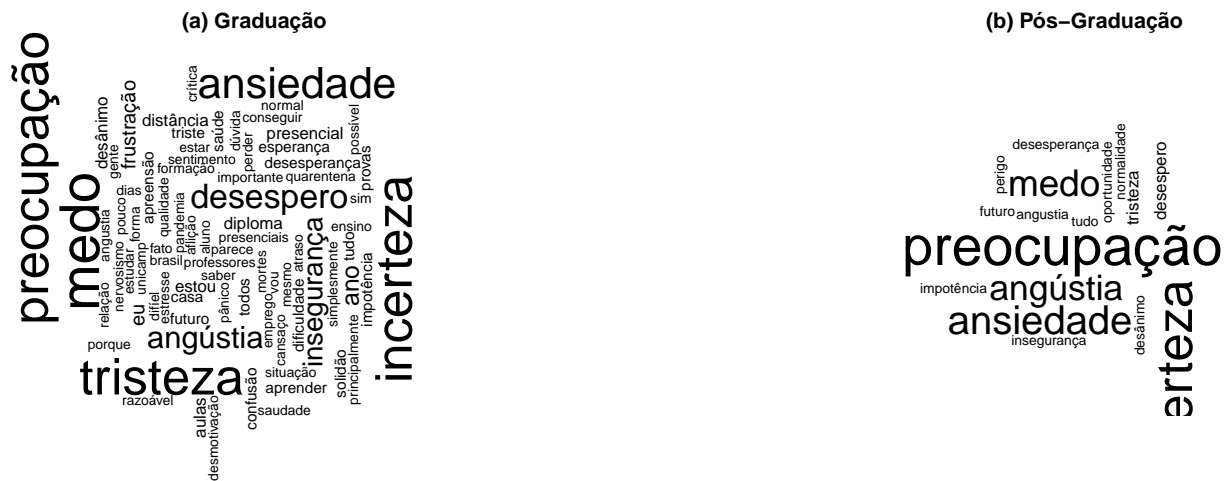


Figura 3: Nuvem de palavras que expressam os sentimentos dos alunos de graduação e pós-graduação do IMECC no início da pandemia por COVID-19 no Brasil.

4.2 Resultados da estimação do modelo de respostas graduais

Para ajustar o modelo de respostas graduais, os itens foram separados em itens que tinham conotação positiva (quanto maior a resposta, maior a habilidade de lidar com a pandemia) e conotação negativa (quanto maior a resposta, menor a habilidade de lidar com a pandemia). Assim, a seguinte separação foi feita:

- Perguntas de natureza negativa:
 - “Enfrenta problemas financeiros?”
 - “Sente falta das atividades presenciais?”
 - “Preocupa-se com sua saúde?”
 - “Preocupa-se com sua família?”
 - “Está sedentário?”
- Perguntas de natureza positiva:
 - “Já se adaptou ao isolamento social?”
 - “Tem conseguido estudar satisfatoriamente?”
 - “Tem se alimentado bem?”
 - “Tem dormido com qualidade?”

O modelo pode ser ajustado de forma restrita ou irrestrita, como descrito na Metodologia. As Tabelas 2 a 5 mostram os valores estimados de $b_{i,j}$ (parâmetro de dificuldade ou Dif. i) e de a_i (parâmetro de discriminação ou Dscrmn).

Comparativamente, os parâmetros de discriminação do modelo de respostas negativas irrestrito são muito mais dispersos que os parâmetros de discriminação do modelo de respostas positivas irrestritos. Isso indica que possivelmente, para o modelo de respostas positivas, seja melhor utilizar o modelo restrito pelo conceito da Navalha de Ockham, que é muito mais simples.

Tabela 2: Estimativas do modelo de resposta gradual restrito para as respostas negativas.

	$b_{1,j}$	$b_{2,j}$	$b_{3,j}$	$b_{4,j}$	a_i
Enfrenta problemas financeiros	-0.38	0.86	2.20	3.94	1.22
Sente falta das atividades presenciais	-3.54	-2.22	-1.09	0.06	1.22
Preocupa se com sua saúde	-4.48	-2.63	-1.43	-0.28	1.22
Preocupa se com sua família	-4.83	-3.20	-2.06	-1.13	1.22
Está sedentário	-2.58	-1.10	-0.08	0.62	1.22

Tabela 3: Estimativas do modelo de resposta gradual irrestrito para as respostas negativas.

	$b_{1,j}$	$b_{2,j}$	$b_{3,j}$	$b_{4,j}$	a_i
Enfrenta problemas financeiros	-0.35	0.69	1.84	3.45	0.26
Sente falta das atividades presenciais	-3.53	-2.25	-1.15	-0.03	1.11
Preocupa se com sua saúde	-9.12	-5.58	-3.11	-0.88	3.66
Preocupa se com sua família	-12.51	-8.95	-5.85	-3.42	5.12
Está sedentário	-2.36	-1.05	-0.15	0.47	0.72

Tabela 4: Estimativas do modelo de resposta gradual restrito para as respostas positivas.

	$b_{1,j}$	$b_{2,j}$	$b_{3,j}$	$b_{4,j}$	a_i
Já se adaptou ao isolamento social	-4.00	-1.92	0.32	2.61	2.01
Tem conseguido estudar satisfatoriamente	-2.74	0.24	2.23	4.00	2.01
Tem se alimentado bem	-4.76	-2.21	-0.18	2.16	2.01
Tem dormido com qualidade	-2.75	-0.65	0.96	2.81	2.01

Tabela 5: Estimativas do modelo de resposta gradual irrestrito para as respostas positivas.

	Dif. 1	Dif. 2	Dif. 3	Dif. 4	Dscrmn
	$(b_{1,j})$	$(b_{2,j})$	$(b_{3,j})$	$(b_{4,j})$	(a_i)
Já se adaptou ao isolamento social	-3.68	-1.76	0.29	2.40	1.71
Tem conseguido estudar satisfatoriamente	-2.72	0.25	2.22	3.97	1.99
Tem se alimentado bem	-5.29	-2.47	-0.20	2.39	2.40
Tem dormido com qualidade	-2.78	-0.66	0.96	2.84	2.05

4.2.1 Diagnóstico

Para decidir entre o modelo restrito e irrestrito, foi realizado o teste ANOVA para o caso de respostas positivas e outro teste para o caso de respostas negativas e os resultados do teste encontram-se na Tabelas 6 e 7, respectivamente. Na Tabela 6 o p-valor <0.001 indica que o modelo irrestrito é melhor que o restrito para o grupo das perguntas negativas. Já a Tabela 7 traz um p-valor = 0.305, o que indica que não temos evidências suficientes para apontar que o modelo restrito é o mais apropriado para ajustar os itens associados a perguntas positivas. Dessa forma, ficaremos com o modelo restrito para o grupo de respostas positivas.

Tabela 6: Resultado do teste ANOVA para comparação do modelo restrito com o irrestrito, segundo perguntas de natureza negativa.

	AIC	Função Ver.	LRT	g.l.	p-valor
Modelo Restrito	3893.34	-1925.67			
Modelo Irrestrito	3758.99	-1854.50	142.35	4	<0.001

Tabela 7: Resultado do teste ANOVA para comparação do modelo restrito com o irrestrito, segundo perguntas de natureza positiva.

	AIC	Função Ver.	LRT	g.l.	p-valor
Modelo Restrito	3427.58	-1696.79			
Modelo Irrestrito	3429.96	-1694.98	3.62	3	0.305

4.2.2 Análises gráficas

Os gráficos das curvas características para as perguntas de natureza negativa encontram-se na Figura 4 e o gráfico de informação dos itens na Figura 5.

A partir da Figura 4, na pergunta sobre ter que lidar com problemas financeiros, temos que a medida que a inabilidade aumenta, a probabilidade do aluno selecionar as opções 1 a 5 cai para aproximadamente 0.2. Quando a inabilidade é mínima a probabilidade do aluno discordar fortemente que enfrenta problemas financeiros é de aproximadamente 0.6, enquanto a probabilidade de selecionar as demais opções é, no máximo, 0.2. À medida que a inabilidade de lidar com a pandemia aumenta, a probabilidade do aluno concordar fortemente que sente falta das atividades presenciais converge para 1, enquanto que as demais opções convergem para 0. Quando a inabilidade de lidar com a pandemia é nula ainda existe a probabilidade de aproximadamente 0.4 do aluno concordar fortemente que sente falta das atividades presenciais, enquanto que a probabilidade fica em torno de 0.2 para as demais opções. À medida que a inabilidade de lidar com a pandemia aumenta, a probabilidade do aluno concordar fortemente que se preocupa com sua saúde converge para 1, enquanto que as demais opções convergem para 0. Quando a inabilidade de lidar com a pandemia é mínima, a probabilidade de discordar que se preocupa com a saúde converge para 1, enquanto que as demais opções tem probabilidade zero de acontecer.

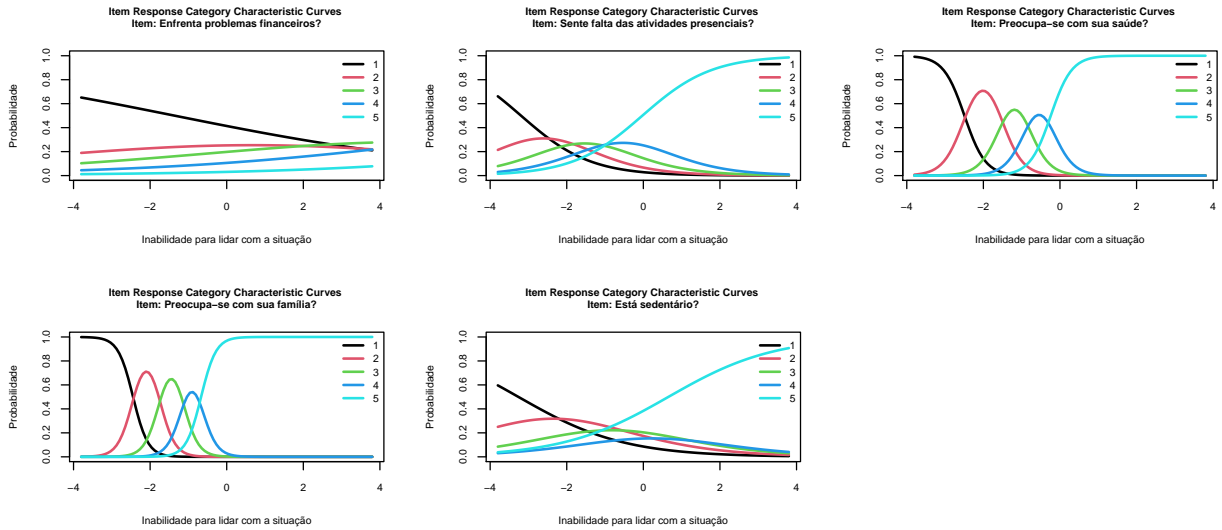


Figura 4: Curvas características para cada categoria das perguntas da classe negativa.

A Figura 5 mostra as curvas de informação de cada item. Vemos que todas os itens avaliam a variável latente muito melhor para níveis mais baixos dela. Também vemos novamente um destaque dos itens *Preocupa-se com sua saúde?* e *Preocupa-se com sua família?*, o que indica que esses itens definem muito melhor a variável latente.

O gráfico da Figura 5 indica que os itens que trazem mais informação sobre a incapacidade de lidar com a pandemia são os itens *Preocupa-se com sua saúde?* e *Preocupa-se com sua família?*. Note que para as incapacidades maiores do que zero os itens não trazem informação.

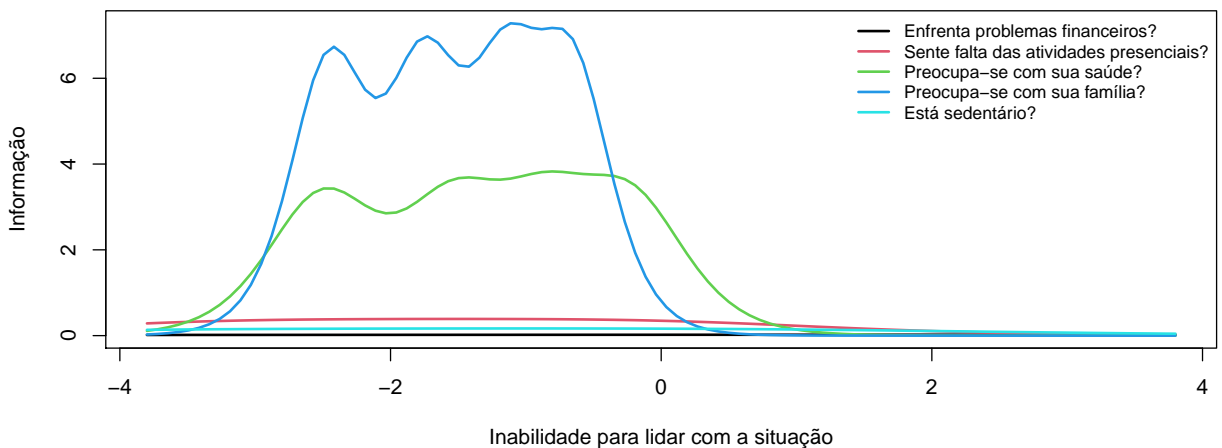


Figura 5: Curvas de informação do item - Negativo

Os gráficos das curvas características para as perguntas de natureza positiva

encontram-se na Figura 6. A curva está deslocada para a esquerda indicando que as respostas dos alunos estão mais associadas a uma inabilidade menor do que zero de lidar com a pandemia. Este comportamento indica que a inabilidade para lidar com a situação é mais informativa quando os alunos reagem mais positivamente diante da pandemia, já que a informação é positiva para valores da inabilidade menores do que zero.

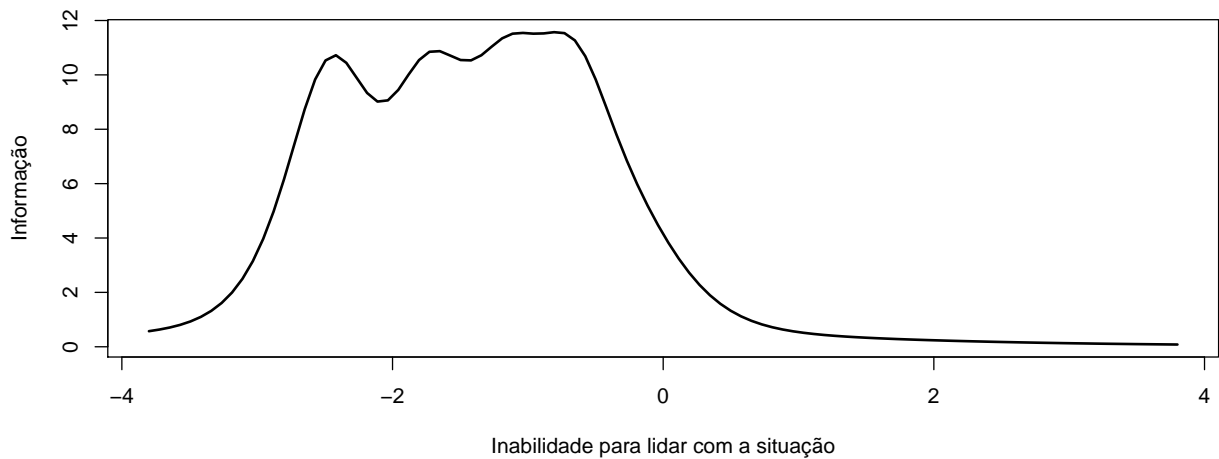


Figura 6: Curva de informação do modelo construído a partir das perguntas da classe negativa.

Os gráficos das curvas características para as perguntas de natureza negativa encontram-se na Figura 7 e o gráfico de informação dos itens na Figura 8.

Com a Figura 7, vemos que enquanto a habilidade de lidar com a pandemia aumenta, a probabilidade do aluno selecionar a opção 1 (Discordo fortemente) vai de aproximadamente 0,9 para 0, enquanto a probabilidade de selecionar a opção 5 cresce de 0 para aproximadamente 0,9. A probabilidade das três opções intermediárias convergem para zero nos extremos. Esse mesmo comportamento é observado nos itens relacionados à satisfação com os estudos, com a alimentação e a qualidade do sono.

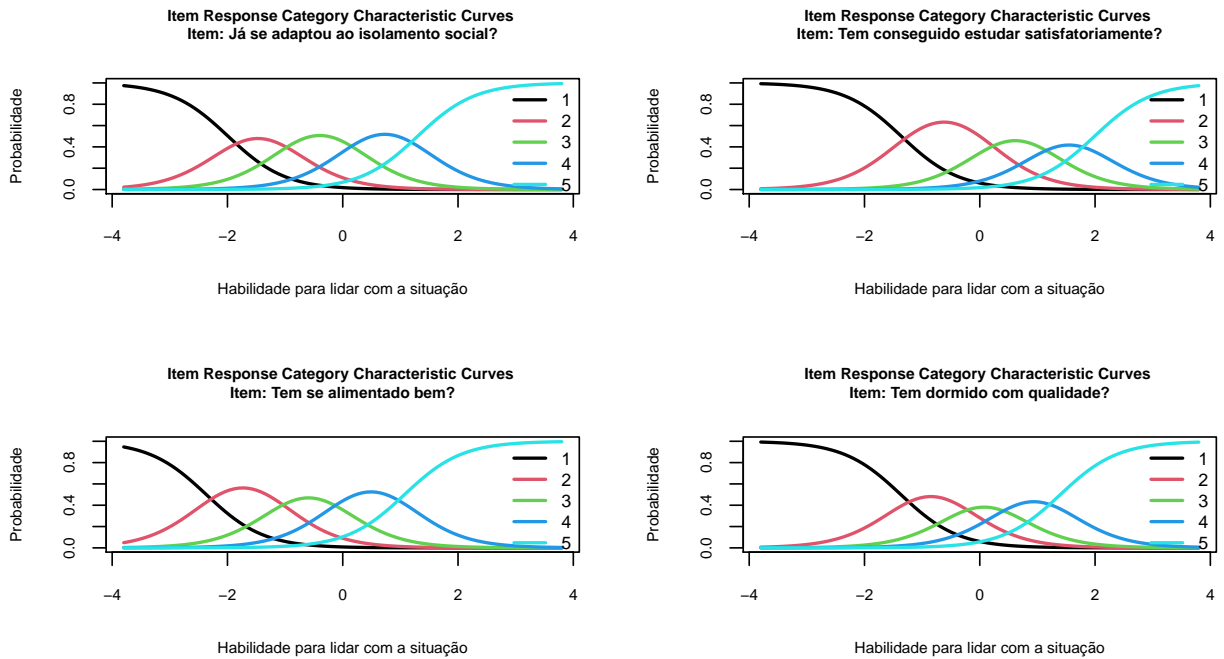


Figura 7: Curvas características para cada categoria das perguntas da classe positiva.

O gráfico da Figura 8 indica que todos os itens trazem informação equivalente sobre a habilidade de lidar com a pandemia, sendo que a informação diminui um pouco nos extremos.

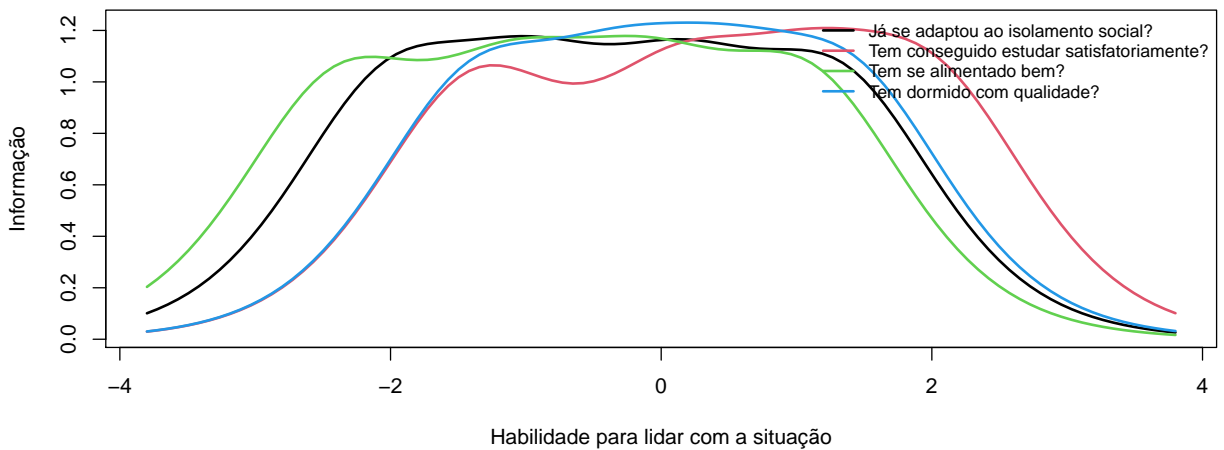


Figura 8: Curvas de informação dos itens classificados como perguntas positivas.

Finalizando, a Figura 9 indica que a habilidade em lidar com a situação é informativa ao longo de todo o traço latente.

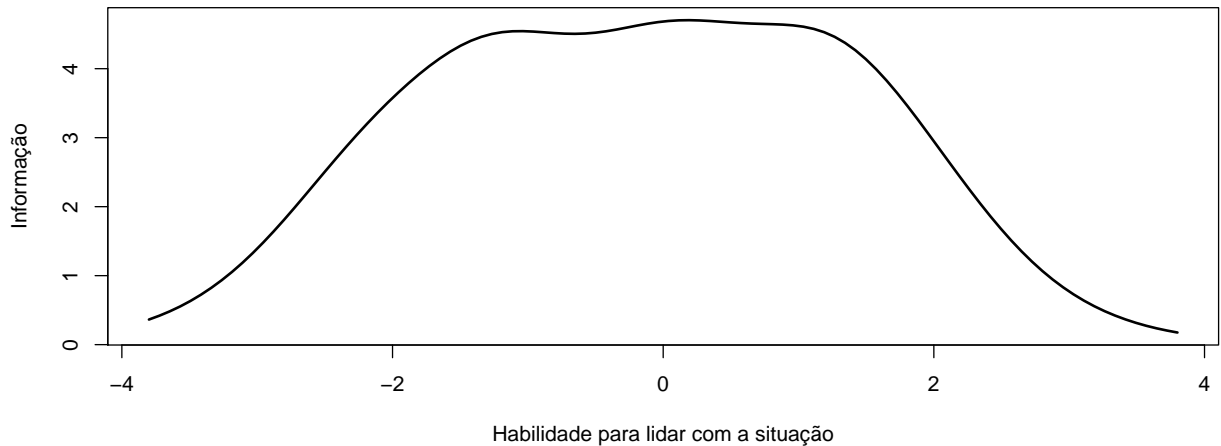


Figura 9: Curva de informação do modelo construído a partir das perguntas da classe positiva.

5 Conclusão

Neste estudo foi possível identificar um cenário dos sentimentos dos alunos quanto aos desfechos da pandemia por Coronavírus em 2020 com reflexos na saúde física e mental destes sujeitos. Os resultados reportados representam os alunos que de forma voluntária responderam ao questionário submetido pela direção do IMECC. O universo de respondentes corresponde a aproximadamente 20% dos alunos do instituto. A frequência relativa ao longo das respostas dadas pelos alunos de graduação foi sempre maior do que às dos alunos de pós-graduação, embora o perfil da distribuição de frequência das categorias seja similar nos dois grupos. Os sentimentos de incerteza, ansiedade, preocupação, medo e angústia são comuns aos dois grupos. Porém, palavras como medo e tristeza são mais frequentes para alunos de graduação, enquanto angústia foi mais frequente entre os alunos de pós-graduação.

Os itens sobre preocupação com a família e com a própria saúde tem categorias facilmente distinguíveis para a definição da variável latente. Esses mesmos itens são os únicos que de fato contribuem para a definição da variável latente. A curva de informação para o modelo ajustado a partir das perguntas de natureza negativa indica que a inabilidade de lidar com a pandemia não é tão ruim, já que a informação está concentrada para os valores negativos da inabilidade.

Todos os itens do grupo de perguntas de natureza positiva tem categorias facilmente distinguíveis para a definição da variável latente. Existe simetria nas respostas dadas pelos alunos respondentes para os itens de natureza positiva, o que indica que existe uma distribuição simétrica da probabilidade de escolha dos itens ao longo da variável latente.

Para um maior alcance e representatividade do efeito da pandemia por coronavírus no alunos do IMECC seria necessário que mais alunos tivessem respondido ao questionário enviado. De qualquer forma, os resultados apontam que a pandemia trouxe sequelas aos alunos do Instituto.

6 Apêndice

6.1 Colunas

- As disciplinas em que você está matriculado estão com atividades online?
- Carimbo de data/hora
- Como está sendo a experiência das aulas remotas?
- E-mail: (opcional, mas preencha caso deseje receber retorno)
- Em média, como você avalia que está sendo seu aprendizado nas disciplinas em está cursando remotamente?
- Enfrenta problemas financeiros?
- Escolha sua categoria.
- Espaço extra para comentar qualquer outra coisa.
- Está sedentário?
- Já se adaptou ao isolamento social?
- Nome: (opcional, mas preencha caso deseje receber retorno)
- Preocupa-se com sua família?
- Preocupa-se com sua saúde?
- Que sentimento primeiro lhe vem à mente em relação ao cenário presente que estamos vivendo?
- Que tipo de dispositivo você utiliza para acessar as aulas/atividades? Marque todos que utilizar.
- Sente falta das atividades presenciais?
- Seu acesso à internet é feito por
- Tem conseguido estudar satisfatoriamente?
- Tem dormido com qualidade?

- Tem se alimentado bem?
- Você solicitou algum recurso a universidade (PRG, IMECC, etc), tal como computador ou chip para acesso à internet, para poder acompanhar as atividades online?

Referências

- H. Akaike. *Information theory and an extension of the maximum likelihood principle*. Akadémiai Kiadó, Armenia, USSR, 1978.
- Thiago Baldini da Silva, Claudia Aparecida Salvuchi, Daniel Kenzi Kanashiro, Flávio Magalhães Braga, Giovanna Maria R. dos Santos, and Daniel Kenzi Kanashiro. Anuário estatístico. *AEPLAN/UNICAMP, Universidade Estadual de Campinas, Assessoria de Economia e Planejamento*. – Campinas, SP, 2020. URL <https://www.aeplan.unicamp.br/anuario/2020/anuario2020.pdf>.
- Dimitris Rizopoulos. ltm: An r package for latent variable modelling and item response theory analyses. *Journal of Statistical Software*, 17(5):1–25, 2006. URL <http://www.jstatsoft.org/v17/i05/>.
- F. Samejima. Estimation of latent ability using a response pattern of graded scores. *Educational testing service, Princeton, New Jersey*, 1968.
- Hadley Wickham, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Golemund, Alex Hayes, Lionel Henry, Jim Hester, Max Kuhn, Thomas Lin Pedersen, Evan Miller, Stephan Milton Bache, Kirill Müller, Jeroen Ooms, David Robinson, Dana Paige Seidel, Vitalie Spinu, Kohske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo, and Hiroaki Yutani. Welcome to the tidyverse. *Journal of Open Source Software*, 4(43):1686, 2019. doi: 10.21105/joss.01686.