

Laboratório 1 - Análise TCT

EST5107 - Modelos de variáveis latentes

Thiago Ferreira Miranda - 11925711

São Carlos, 14/08/2023

Resumo

Este trabalho busca consolidar conhecimentos da Teoria Clássica dos Testes (TCT) analisando itens de três testes: a prova de matemática do Exame Nacional do Ensino Médio (ENEM) do ano de 2022; a prova matemática do CRESCER de 1998; e o estudo comparativo de professores espanhóis e peruanos de educação primária em relação à Estatística.

1 Introdução

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 indivíduos, 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, a Teoria Clássica dos Testes (TCT) utiliza foi pioneira ao apresentar, além da avaliação via pontuação total do teste ou escore total, algumas medidas para avaliar a qualidade do instrumento de medida.

Este trabalho tem como objetivo analisar três instrumentos de medida por meio da lente da Teoria Clássica dos Testes. Ao empregar essa abordagem, pretende-se examinar a qualidade e a consistência dos instrumentos escolhidos.

2 Métodos

2.1 Exame Nacional do Ensino Médio

O Exame Nacional do Ensino Médio (ENEM) foi criado pelo Ministério da Educação (MEC), onde sua primeira aplicação ocorreu em 1998. A implementação do referido exame teve como um dos objetivos avaliar o desempenho dos alunos após a conclusão do Ensino Médio, para poder contribuir na melhoria da educação do país. E atualmente, o ENEM tem sido um importante mecanismo na implementação da reforma do Ensino Médio. [4]

No início de sua aplicação, em 1998, o ENEM utilizava o método da Teoria Clássica dos Testes (TCT) para realizar a avaliação da competência dos estudantes em relação às áreas de conhecimento aplicados no referido exame. A TCT teve como objetivo verificar a habilidade do estudante, no qual a análise da competência e a interpretação dos dados se deram com os resultados relacionados à prova como um todo. [5]

Com o passar dos anos, o ENEM passou a ser utilizado no critério da seleção de bolsas no Programa Universidade para Todos (Prouni) e, logo em seguida, os resultados obtidos pelo exame passaram a ser utilizados como forma de seleção para o ingresso no ensino superior em várias universidades do país (TAVARES, 2010). Por conta disso, o ENEM reformulou sua metodologia, passando a utilizar a Teoria de Resposta ao Item (TRI). Após algumas

mudanças, o ENEM passou a ter 180 questões, juntamente com uma redação, no qual as questões são divididas em quatro áreas do conhecimento: Linguagens, Códigos e suas Tecnologias, Matemática e suas Tecnologias, Ciências da Natureza e suas Tecnologias e Ciências Humanas e suas Tecnologias.

Este trabalho realiza a análise via TCT apenas da área de Matemática e suas Tecnologias para a aplicação regular do ano de 2022.

2.2 Teoria clássica dos testes

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 indivíduos participantes dos grupos observados. 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.

2.2.1 Índice de dificuldade

Proporção de acertos do itens, valor que varia de 0 a 1. Espera-se que o instrumento de medida seja composto por itens com dificuldades variadas.

$$Diff = \frac{\sum_{i=1}^N I_i}{N} \quad (1)$$

em que,

I_i é a variável indicadora que assume 1 quando o item é respondido corretamente pelo respondente i e 0, caso contrário;

N é a quantidade total de respondentes;

$\sum_{i=1}^N I_i$ é a soma de resposta corretas no item.

2.2.2 Índice de discriminação

Diferença entre a proporção de acerto do grupo de 33% dos respondentes com os escores mais altos e a proporção de acerto do grupo de 33% dos respondentes com os escores mais baixos. Com valores podendo variar de -1 a 1. Espera-se que este índice seja positivo, aferindo que respondentes com escores mais altos tem maior facilidade com o item que respondentes com escores mais baixos.

2.2.3 Coeficiente de correlação ponto-bisserial

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

$$\rho_{pb} = \frac{\bar{X}_A - \bar{X}_T}{S_T} \sqrt{\frac{p}{1-p}} \quad (2)$$

em que,

\bar{X}_A é a média dos escores dos respondentes que acertaram o item;

\bar{X}_T é a média global dos escores do teste;

S_T é o desvio padrão do teste;

p é a proporção de respondentes que acertaram o item.

2.2.4 Coeficiente de correlação bisserial

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

$$\rho_b = \rho_{pb} \frac{\sqrt{p(1-p)}}{h(p)} \quad (3)$$

em que,

ρ_{pb} é a correlação ponto-bisserial;

p é a proporção de respondentes que acertaram o item;

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

2.2.5 Coeficiente alfa de Cronbach

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

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum s_i^2}{s_T^2} \right) \quad (4)$$

em que,

n é o número de itens;

$\sum s_i^2$ é soma das variâncias dos n itens;

s_T^2 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.

3 Resultados

Tabela 1: Estatística dos itens de Matemática do ENEM 2022 via teoria clássica dos testes.

Item	Dificuldade	Discriminação	Desvio-Padiao	Correlação Ponto-Bisserial	Correlação Ponto-Bisserial Sem	Correlação Bisserial	Correlação Bisserial Sem	Alpha Sem
005961	0.30	0.24	0.46	0.25	0.17	0.33	0.22	0.72
008364	0.30	0.23	0.46	0.24	0.16	0.32	0.21	0.72
010322	0.49	0.46	0.50	0.40	0.31	0.50	0.39	0.71
010409	0.39	0.32	0.49	0.30	0.21	0.38	0.27	0.72
010500	0.19	0.09	0.39	0.14	0.06	0.20	0.09	0.72
014797	0.15	0.11	0.36	0.23	0.16	0.35	0.25	0.72
019807	0.09	0.04	0.28	0.22	0.17	0.39	0.30	0.72
028683	0.26	0.12	0.44	0.21	0.13	0.29	0.18	0.72
030053	0.31	0.24	0.46	0.29	0.21	0.38	0.28	0.72
030327	0.56	0.57	0.50	0.47	0.39	0.59	0.49	0.71
030493	0.30	0.25	0.46	0.25	0.17	0.33	0.22	0.72
031516	0.31	0.31	0.46	0.34	0.26	0.44	0.33	0.71
032369	0.20	0.12	0.40	0.24	0.16	0.34	0.24	0.72
043182	0.34	0.38	0.47	0.42	0.34	0.54	0.44	0.71
047309	0.32	0.26	0.47	0.26	0.18	0.34	0.23	0.72
053461	0.25	0.17	0.43	0.24	0.16	0.33	0.22	0.72
060441	0.67	0.44	0.47	0.37	0.29	0.49	0.38	0.71
061248	0.24	0.21	0.43	0.29	0.21	0.40	0.29	0.72
063204	0.23	0.27	0.42	0.39	0.32	0.54	0.44	0.71
063646	0.29	0.23	0.45	0.24	0.16	0.32	0.21	0.72
068123	0.29	0.49	0.46	0.51	0.45	0.68	0.59	0.70
068369	0.27	0.40	0.44	0.46	0.39	0.62	0.53	0.71
075829	0.29	0.28	0.45	0.30	0.22	0.39	0.29	0.72
082581	0.18	0.19	0.38	0.31	0.24	0.45	0.35	0.72
085013	0.23	0.16	0.42	0.21	0.14	0.30	0.19	0.72
085228	0.24	0.09	0.43	0.10	0.02	0.13	0.02	0.73
085343	0.15	0.09	0.35	0.14	0.08	0.22	0.12	0.72
085588	0.62	0.47	0.49	0.40	0.32	0.51	0.41	0.71
086466	0.24	0.12	0.43	0.21	0.13	0.29	0.18	0.72
086499	0.21	0.07	0.41	0.08	0.01	0.12	0.01	0.73
086840	0.29	0.18	0.45	0.16	0.08	0.21	0.10	0.72
089637	0.09	0.02	0.29	0.02	-0.03	0.04	-0.05	0.73
095509	0.16	0.13	0.36	0.24	0.18	0.37	0.27	0.72
095676	0.54	0.52	0.50	0.43	0.35	0.54	0.44	0.71
096315	0.49	0.47	0.50	0.40	0.32	0.51	0.40	0.71
097590	0.18	0.15	0.39	0.26	0.19	0.38	0.28	0.72
111516	0.35	0.41	0.48	0.42	0.34	0.54	0.44	0.71
111738	0.30	0.18	0.46	0.24	0.16	0.32	0.21	0.72
117651	0.27	0.23	0.44	0.23	0.15	0.31	0.20	0.72
117742	0.35	0.23	0.48	0.22	0.13	0.28	0.17	0.72
117820	0.26	0.16	0.44	0.16	0.08	0.21	0.10	0.72
117877	0.33	0.37	0.47	0.40	0.33	0.52	0.42	0.71
117886	0.19	0.11	0.39	0.12	0.04	0.17	0.06	0.72
117973	0.23	0.10	0.42	0.15	0.07	0.21	0.10	0.72

Tabela 2: Estatística dos itens de Matemática CRESCER 1998 via teoria clássica dos testes.

Item	Dificuldade	Discriminação	Desvio-Padiao	Correlação Ponto-Bisserial	Correlação Ponto-Bisserial Sem	Correlação Bisserial	Correlação Bisserial Sem	Alpha Sem
i01	0.79	0.37	0.41	0.35	0.14	0.50	0.20	0.47
i02	0.85	0.23	0.35	0.25	0.06	0.38	0.09	0.49
i03	0.52	0.56	0.50	0.49	0.24	0.61	0.30	0.44
i04	0.93	0.16	0.25	0.38	0.26	0.73	0.49	0.45
i05	0.87	0.23	0.34	0.32	0.15	0.51	0.23	0.47
i06	0.37	0.47	0.48	0.39	0.14	0.50	0.18	0.48
i07	0.92	0.16	0.27	0.36	0.23	0.67	0.42	0.45
i08	0.88	0.30	0.33	0.48	0.33	0.78	0.53	0.43
i09	0.79	0.21	0.41	0.26	0.04	0.36	0.05	0.50
i10	0.86	0.26	0.35	0.35	0.17	0.54	0.27	0.46
i11	0.93	0.21	0.25	0.50	0.38	0.95	0.73	0.43
i12	0.35	0.47	0.48	0.43	0.18	0.55	0.23	0.46
i13	0.82	0.28	0.38	0.36	0.16	0.53	0.24	0.46
i14	0.95	0.05	0.23	0.18	0.06	0.37	0.12	0.48

Tabela 3: Estatística dos itens de Atitudes em Estatística via teoria clássica dos testes.

Item	Mean.Score	Item.Difficulty	Point.Biserial	Polyserial	Prop.1	Prop.2	Prop.3	Prop.4	Prop.5	Point.Biserial.1	Point.Biserial.2	Point.Biserial.3	Point.Biserial.4	Point.Biserial.5
Item 1	3.29	0.66	0.12	0.13	0.03	0.25	0.26	0.34	0.12	0.02	0.01	-0.17	-0.01	0.22
Item 2	3.99	0.80	0.36	0.43	0.02	0.05	0.05	0.68	0.20	-0.25	-0.17	-0.10	0.02	0.21
Item 3	3.81	0.76	0.30	0.33	0.01	0.07	0.18	0.56	0.17	-0.06	-0.08	-0.22	0.02	0.28
Item 4	3.53	0.71	0.28	0.32	0.04	0.10	0.20	0.60	0.05	-0.33	0.01	-0.08	0.12	0.15
Item 5	4.14	0.83	0.37	0.42	0.03	0.05	0.11	0.39	0.42	-0.02	-0.28	-0.21	-0.07	0.33
Item 6	2.99	0.60	0.34	0.36	0.09	0.21	0.38	0.27	0.05	-0.16	-0.29	0.08	0.23	0.09
Item 7	3.10	0.62	0.60	0.64	0.10	0.34	0.08	0.34	0.15	-0.28	-0.40	0.00	0.29	0.38
Item 8	3.55	0.71	0.31	0.35	0.01	0.21	0.13	0.54	0.12	-0.05	-0.21	-0.07	0.05	0.28
Item 9	3.56	0.71	0.46	0.50	0.04	0.11	0.24	0.47	0.14	-0.13	-0.24	-0.28	0.20	0.34
Item 10	3.44	0.69	0.46	0.49	0.01	0.18	0.27	0.44	0.10	-0.03	-0.25	-0.29	0.25	0.34
Item 11	3.77	0.75	0.56	0.61	0.01	0.10	0.16	0.56	0.16	-0.19	-0.39	-0.09	0.01	0.44
Item 12	3.58	0.72	0.56	0.60	0.03	0.17	0.16	0.47	0.17	-0.31	-0.26	-0.20	0.13	0.42
Item 13	3.19	0.64	0.29	0.31	0.06	0.27	0.21	0.34	0.12	-0.18	-0.19	0.03	0.15	0.13
Item 14	3.71	0.74	0.24	0.28	0.01	0.11	0.13	0.64	0.10	-0.23	-0.15	-0.03	0.15	0.05
Item 15	3.08	0.62	0.37	0.39	0.07	0.26	0.29	0.28	0.10	-0.20	-0.21	0.01	0.15	0.24
Item 16	3.03	0.61	0.52	0.55	0.11	0.33	0.15	0.25	0.16	-0.16	-0.40	0.00	0.22	0.39
Item 17	3.90	0.78	0.45	0.51	0.04	0.11	0.05	0.50	0.29	-0.04	-0.31	-0.23	-0.12	0.47
Item 18	3.56	0.71	0.14	0.15	0.02	0.08	0.40	0.32	0.18	-0.23	-0.13	0.22	-0.25	0.20
Item 19	3.60	0.72	0.58	0.62	0.05	0.19	0.15	0.33	0.28	-0.27	-0.35	-0.13	0.06	0.47
Item 20	3.51	0.70	0.66	0.71	0.08	0.15	0.18	0.37	0.22	-0.45	-0.30	-0.12	0.18	0.44
Item 21	3.79	0.76	0.34	0.38	0.01	0.10	0.15	0.54	0.19	-0.13	-0.18	-0.07	-0.07	0.34
Item 22	3.83	0.77	0.56	0.61	0.03	0.08	0.16	0.46	0.26	-0.30	-0.33	-0.19	0.13	0.35

4 Discussão de resultados

5 Referências

References

- [1] ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. *Teoria da resposta ao item: conceitos e aplicações*. São Paulo : ABE - Associação Brasileira de Estatística, 2000.
- [2] ANJOS, A.; ANDRADE, D. F. *Teoria de Resposta ao Item com o uso do R*. In: Simpósio Nacional de Probabilidade e Estatística, 2012.
- [3] R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. 2023, URL <https://www.R-project.org/>.
- [4] CASTRO, M. H. G.; TIEZZI, S. *A reforma do ensino médio e a implementação do Enem no Brasil*. In: *Os desafios da educação no Brasil*. 1. ed. Rio de Janeiro: Nova Fronteira, 2005. p. 113-140.
- [5] PASQUALI, L. *Psicometria*. Revista da Escola de Enfermagem da USP, v.43, p. 993-994,2009.
- [6] TAVARES, H. R. *A Prova dos Nove da Educação Brasileira*. Jornal Beira do Rio. Belém, PA, 2010. n. 130. Nov/2010.