POLITANO APROVA

Como a TRI constrói a sua nota

Matemática no ENEM: por que dois alunos com os mesmos acertos recebem notas diferentes - e o que isso muda na sua preparação

Prof. Yuri Politano

Material gratuito · politanoaprova.com.br · 2026

Apresentação

Este material nasceu de um levantamento que fizemos aqui no Politano Aprova com as provas de Matemática do ENEM de 2015 a 2025: classificamos as 990 questões por assunto, medimos a dificuldade de cada uma e cruzamos o número de acertos com as notas reais registradas nos microdados oficiais do INEP.

Ao olhar essas tabelas, uma pergunta aparece sempre: como pode um aluno com 33 acertos tirar nota maior do que outro, no mesmo ano, com os mesmos 33 acertos? A resposta está na Teoria de Resposta ao Item, a TRI, o método que o ENEM usa desde 2009 para transformar o seu cartão de respostas em nota.

Este texto foi escrito para alunos. A ideia é que, ao final, você saiba:

- por que a nota do ENEM não é a soma dos acertos;
- o que são os três parâmetros de cada questão;
- o que é o tal do padrão de respostas (e o que é mito sobre ele);
- o que significa uma nota 500, 700 ou acima de 1.000;
- e como o cálculo acontece de verdade, para quem quiser ver a matemática por dentro.

Tudo aqui tem como base documentos oficiais do INEP, listados no final. Quando usarmos números para ilustrar, avisaremos que são exemplos fictícios, criados por nós só para mostrar a lógica do cálculo.

1. A nota do ENEM não é a soma dos acertos

Na escola, a prova costuma valer 10 pontos: cada questão vale uma fatia, e a nota é a soma das fatias. Esse jeito de medir tem um nome: Teoria Clássica dos Testes. Ele funciona bem dentro de uma sala de aula, mas tropeça quando se quer comparar pessoas que fizeram provas diferentes.

Pense: se a prova de 2024 foi um pouco mais fácil que a de 2023, 30 acertos em 2024 valem o mesmo que 30 acertos em 2023? Pela soma simples, sim. Na prática, não. E o ENEM precisa que valham o mesmo, porque as notas alimentam Sisu, ProUni e Fies ano após ano, com provas diferentes a cada edição.

Foi para resolver isso que o INEP adotou a TRI em 2009. A ideia central é trocar a pergunta "quantas questões você acertou?" por "qual é o seu nível de proficiência em Matemática?". A proficiência é medida numa régua única, a escala do ENEM, e cada questão ocupa um lugar conhecido nessa régua. Provas diferentes, montadas com questões diferentes, passam a medir na mesma régua - e os resultados ficam comparáveis entre anos.

Uma consequência importante: anular uma questão não quebra o cálculo. Como cada questão tem seu lugar na escala, o modelo estima a sua proficiência a partir das questões restantes. O próprio INEP afirma que a anulação de um item, ou mesmo a aplicação de provas com itens diferentes, não compromete a estimação da nota.

> **Fique atento**
> Na TRI, não basta contar quantas questões você acertou. Importa quais questões você acertou - e em que combinação elas aparecem no seu padrão de respostas.

2. Cada questão tem três parâmetros

No modelo usado pelo ENEM, cada questão é descrita por três números, chamados parâmetros do item. Eles são estimados em pré-testes, antes de a questão entrar no banco de itens do INEP.

O parâmetro b: a dificuldade

É o mais intuitivo. O b indica em que ponto da escala de proficiência a questão está posicionada: a "altura" dela na régua. Uma questão fácil tem b baixo (digamos, na altura de 350); uma difícil tem b alto (digamos, na altura de 680). Quem tem proficiência bem acima do b de uma questão tem alta probabilidade de acertá-la; quem está bem abaixo tende a errar.

O parâmetro a: a discriminação

Mede o quanto a questão separa quem sabe de quem não sabe. Em uma questão com a alto, a probabilidade de acerto sobe rapidamente quando a proficiência passa do ponto b: ela distingue bem alunos de níveis próximos. Em uma questão com a baixo, a curva sobe devagar: acertá-la ou errá-la diz pouco sobre o seu nível.

O parâmetro c: o acerto casual

Em prova de múltipla escolha, quem não domina o conteúdo ainda pode acertar no chute. O parâmetro c representa exatamente isso: a probabilidade de acerto casual daquela questão. Ele é estimado para cada item - não é um valor fixo nem igual para todas. É por causa dele que o modelo consegue olhar para um acerto "fora da curva" e tratá-lo como possível chute, como veremos na próxima seção.

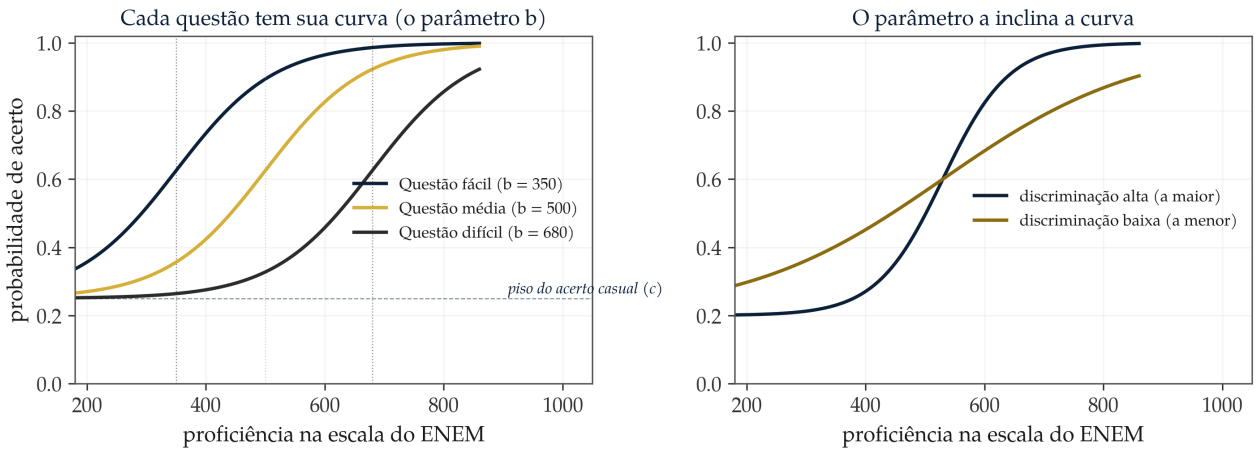


Figura 1 - A curva característica de cada questão. À esquerda, três questões com dificuldades (b) diferentes; a linha tracejada mostra o piso do acerto casual (c). À direita, o efeito da discriminação (a) na inclinação.

Parâmetro não é peso
Circula por aí a ideia de que as questões têm "pesos" secretos e que, conhecendo-os, daria para recalcular a nota com uma média ponderada. O próprio INEP afirma que essa interpretação está errada: os parâmetros alimentam um modelo estatístico completo, não uma conta de soma e multiplicação.

3. O padrão de respostas: por que os mesmos acertos dão notas diferentes

Aqui mora o coração da TRI - e a maior concentração de mitos por metro quadrado.

O modelo parte de um pressuposto pedagógico: o conhecimento é cumulativo. Quem domina habilidades complexas costuma dominar também as mais simples. Então, se um aluno acerta as questões difíceis e erra as fáceis, o padrão dele é incoerente com o que o modelo espera. A explicação mais provável, estatisticamente, é que parte dos acertos difíceis tenha sido casual.

O guia oficial do INEP ilustra com dois participantes que acertaram 5 de 10 questões. O participante A acertou as mais fáceis e errou a partir de certo nível de dificuldade, exatamente como se espera de quem tem proficiência intermediária: nota maior. O participante B acertou as mais difíceis e errou as mais fáceis: padrão incoerente, nota menor. O guia usa a imagem de uma corrida de barreiras: quem salta 80 centímetros no treino deveria saltar com folga as barreiras mais baixas. Se não salta, algo está estranho.

Reproduzimos essa lógica em uma simulação ilustrativa: inventamos uma mini-prova de 10 questões, com parâmetros escolhidos por nós, e calculamos as notas com o mesmo tipo de modelo do ENEM (o logístico de três parâmetros, com o método EAP - a matemática aparece na seção 5). Os dois participantes acertaram 5 questões:

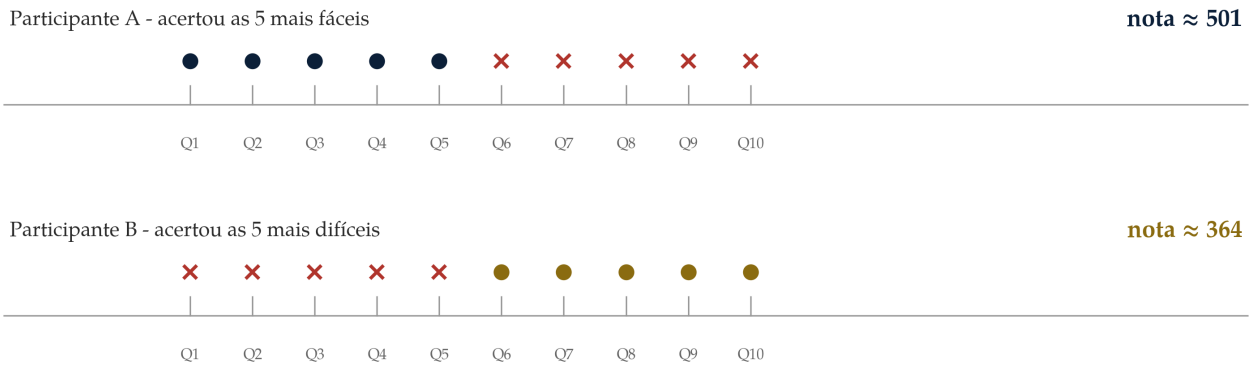


Figura 2 - Mesmo número de acertos (5 em 10), padrões diferentes. Simulação ilustrativa com parâmetros fictícios, usando o modelo 3PL e o método EAP.

Mesmos 5 acertos. Notas a mais de 130 pontos de distância. É por isso que nenhuma tabela de "acertos vezes nota" consegue dar um valor único: ela mostra, no máximo, a faixa em que as notas realmente caíram.

A conta, em versão miniatura

A seção 5 mostra a fórmula completa, que pesa todos os níveis de proficiência de uma vez. Mas a lógica cabe numa conta pequena, com só dois níveis disputando para explicar um cartão: 350 e 550.

Pegue o cartão do Participante B: 5 acertos, todos nas questões difíceis. O modelo combina duas informações sobre cada nível candidato:

- **Chance inicial.** Na distribuição inicial que escolhemos para este exemplo fictício, o nível 550 recebe peso cerca de 2,7 vezes maior que o nível 350. É a "a priori" da seção 5, aqui reduzida a dois níveis: simplificação didática de uma distribuição contínua, não a chance literal de pontos isolados.
- **Compatibilidade com o cartão.** Com os parâmetros da nossa mini-prova, a chance de esse padrão de acertos e erros sair de um aluno de nível 350 é cerca de 39 vezes a chance de sair de um aluno de nível 550. Faz sentido: para quem está em 550, errar as cinco questões fáceis é muito estranho.

Multiplicando as duas informações: o nível 350 fica com peso $1 \times 39 = 39$, e o nível 550, com $2{,}7 \times 1 = 2{,}7$. A nota é a média ponderada pelos pesos finais:

$$(39 \times 350 + 2{,}7 \times 550) \div (39 + 2{,}7) = 15.135 \div 41{,}7 \approx 363$$

O padrão incoerente puxou a nota para perto do nível mais baixo: a compatibilidade com o cartão venceu a chance inicial. A conta completa da seção 5, que pesa todos os níveis, dá 364 para esse mesmo cartão.

Repare no que não apareceu: nenhum ponto por questão, nenhum desconto. Só chances iniciais, compatibilidade e uma média ponderada. (*Versão de brinquedo: o cálculo completo usa todos os níveis, não só dois.*)

O que é mito nessa história

Da versão mal contada dessa história nasceu a lenda de que "errar questão fácil derruba a nota" ou que "a TRI pune quem acerta questão difícil". Não é assim. Não existe penalidade automática nem regra isolada: o que existe é um modelo que olha o padrão inteiro de respostas e estima a proficiência que o torna mais provável. Acertar uma questão difícil nunca diminui a sua nota - no máximo, aquele acerto vale menos do que valeria se viesse acompanhado dos acertos fáceis esperados.

E o próprio guia do INEP é direto: acertar "no chute" não faz a nota diminuir; o acerto simplesmente não vale tanto quanto valeria com coerência. Por isso, a orientação oficial é clara: sempre é melhor responder do que deixar em branco, porque uma questão certa sempre aumenta a nota - e questão em branco é corrigida como errada. Aumentar, aqui, não é somar uma quantia fixa de pontos: é a proficiência estimada que sobe.

4. A escala: o que significam 500, 700 ou mais de 1.000

Quando a TRI foi adotada, em 2009, o INEP precisou fixar uma régua. A escolha: a proficiência média dos concluintes regulares do ensino médio daquele ano foi definida como 500, e o desvio-padrão do grupo, como 100. Esses números são apenas a âncora da régua, como o metro de platina guardado em Paris: uma referência fixa para que todas as edições seguintes possam ser comparadas.

Daí saem duas consequências que respondem a dúvidas clássicas.

Primeira: a sua nota não depende dos outros participantes. Não é "a prova estava fácil, então todo mundo ganhou ponto" nem "a média nacional caiu, então a minha sobe". O valor 500 dos concluintes regulares de 2009 serve só como referência da escala; a sua nota depende do seu padrão de respostas e dos parâmetros das questões da sua edição.

Segunda: a escala não tem mínimo e máximo fixos. Eles dependem das questões de cada prova. Quem erra tudo recebe o valor mínimo que aquela edição consegue medir - e não uma nota zero calculada, porque não dá para concluir, a partir de um teste, que alguém sabe zero. Pela mesma lógica, acertar todas as questões não garante 1.000: garante o máximo que aquela prova mede. E o teto pode passar de 1.000: no levantamento das edições de 2015 a 2025 que acompanha este material, a maior nota de Matemática encontrada nos microdados oficiais foi 1.008,3, no ENEM 2015.

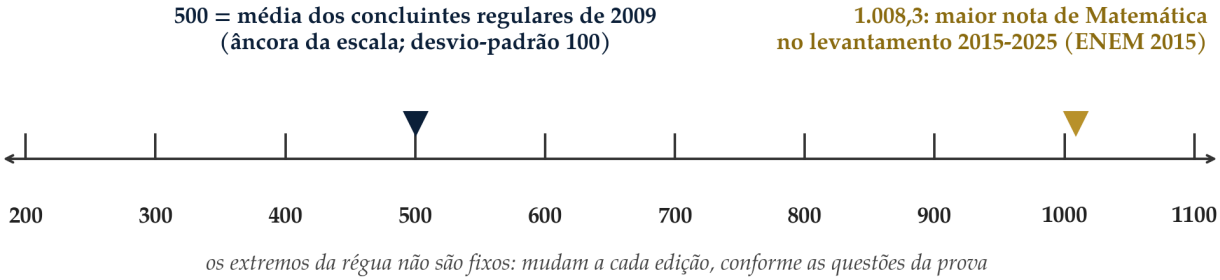


Figura 3 - A escala do ENEM: referência fixa em 500, extremos que variam conforme as questões de cada edição.

Cada área tem a sua régua
As quatro áreas do ENEM têm escalas independentes, ancoradas da mesma forma mas separadas. Comparar 700 em Matemática com 700 em Ciências Humanas não faz sentido: é como comparar quem corre 100 metros com quem nada 100 metros, na analogia do próprio INEP. A redação, aliás, nem usa TRI: ela vale de 0 a 1.000, corrigida por dois avaliadores independentes, com um terceiro em caso de discrepância.

E a nota 0 dos microdados?
Nos microdados oficiais, existe nota 0 registrada. No levantamento que acompanha este material, conferimos esses casos item a item nos anos de 2018, 2019 e 2024: em todos eles, a nota 0 correspondia a folha de respostas totalmente em branco. Esse zero registrado não deve ser confundido com o mínimo da escala estimado pelo modelo.

5. Como o cálculo acontece de verdade

Esta seção é para quem quer ver a engrenagem. Se a matemática pesar, pule para a seção 6 sem medo: o essencial já foi dito. Aqui vamos apresentar o modelo oficial e uma simulação numérica completa, calculada por nós com parâmetros fictícios e uma prior escolhida por nós. Os parâmetros dos itens são públicos: constam dos microdados oficiais do INEP (é de lá que vem o parâmetro b usado no levantamento que acompanha este material). Ainda assim, sem

todos os insumos oficiais usados no processamento de uma edição, nenhum cálculo externo reproduz exatamente a nota oficial - a nossa simulação ilustra o método, não os números de uma edição real.

O modelo: logístico de três parâmetros

Para cada questão, a probabilidade de um participante de proficiência θ acertá-la é dada pela função logística de três parâmetros, proposta por Birnbaum em 1968:

$$P(u_{ji}=1 \mid \theta_j, a_i, b_i, c_i) = c_i + \frac{1-c_i}{1+e^{-a_i(\theta_j - b_i)}}$$

Nela, a, b e c são os três parâmetros da questão que vimos na seção 2 (o índice i só indica de qual questão estamos falando). Quanto maior a proficiência θ, maior a probabilidade de acerto - e a curva nunca desce abaixo de c, o piso do acerto casual.

Uma ressalva de métrica: nas fórmulas desta seção, θ e b aparecem na escala-padrão do modelo (média 0 e desvio 1), e a discriminação a é medida nessa unidade. Para a régua do boletim vale a conversão da equação adiante: multiplicar por 100 e somar 500. Um b de 1,8 na escala-padrão é o mesmo b "na altura de 680" dos exemplos anteriores.

O padrão inteiro entra na conta

Com o cartão de respostas na mão, o modelo pergunta: qual proficiência torna mais provável exatamente este padrão de acertos e erros? Para isso, monta-se a função de verossimilhança, que multiplica as probabilidades de cada resposta observada:

$$L(u_j \mid \theta) = \prod_{i=1}^{n} P_i(\theta)^{u_{ji}} \, [\,1 - P_i(\theta)\,]^{1-u_{ji}}$$

Nela, u vale 1 quando o participante acerta a questão e 0 quando erra. Cada valor possível de proficiência recebe um "peso de plausibilidade" dado por essa função.

A nota é uma média ponderada de proficiências: o método EAP

O ENEM não escolhe simplesmente a proficiência mais provável. Ele usa o método Expected a Posteriori (EAP), o mesmo do Saeb: calcula a média de todos os valores possíveis de proficiência θ, ponderada pela verossimilhança do padrão de respostas e por uma distribuição a priori, que carrega a média e o desvio-padrão dos concluintes do ensino médio da rede pública que fizeram o ENEM 2009 - grupo diferente da referência da escala, que são os concluintes regulares:

$$\hat{\theta}_j^{\mathrm{EAP}} = E(\theta \mid u_j) = \frac{\int \theta \, L(u_j \mid \theta) \, f(\theta) \, d\theta}{\int L(u_j \mid \theta) \, f(\theta) \, d\theta}$$

A integral é resolvida numericamente, por um método chamado quadratura gaussiana. No fim, a proficiência θ é convertida para a escala que aparece no boletim:

$$\text{Nota} = 500 + 100 \times \theta$$

Uma salvaguarda de qualidade digna de nota consta da nota técnica de implantação do método, de 2011: nela, o INEP determinava que o cálculo fosse refeito de forma independente por três equipes - especialistas do Cespe/UnB, da Cesgranrio e do próprio INEP - e que o resultado só fosse divulgado com 100% de concordância entre os três grupos, participante a participante. Não há confirmação pública recente de que esse arranjo siga exatamente igual; fica o registro do cuidado documentado na implantação.

O exemplo ilustrativo, com números

Voltamos à mini-prova fictícia da seção 3: 10 questões, com dificuldades b espaçadas de 300 a 750 na escala, discriminações entre 1,0 e 1,4 e acertos casuais entre 12% e 25%. Os dois participantes acertaram 5 questões cada. Aplicando o modelo acima, obtemos a distribuição a posteriori de cada padrão:

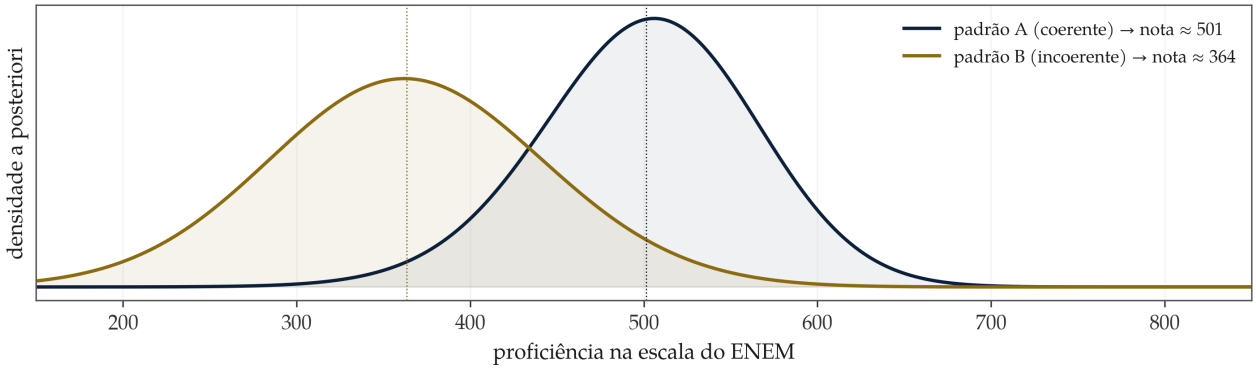


Figura 4 - A "paisagem" de proficiências plausíveis para cada padrão de respostas. A nota EAP é o centro de gravidade de cada curva. Exemplo ilustrativo, com parâmetros fictícios.

O participante A, coerente, fica com nota aproximada de 501. O participante B, que acertou as difíceis mas errou as fáceis, fica com 364: o modelo explica boa parte dos seus acertos difíceis como casuais, porque a base que deveria sustentá-los não aparece. Um terceiro padrão, com 5 acertos alternados, cai entre os dois, em 451. Mesma prova, mesmo número de acertos, três notas.

Por que não existe fórmula de acertos para nota
A nota depende de quais questões você acertou, dos três parâmetros de cada uma e do padrão completo de respostas. Nenhuma tabela, planilha ou simulador reproduz isso por fora: todos são aproximações. O que dá para fazer com honestidade - e é o que fazemos no levantamento irmão deste material - é olhar para os resultados reais de cada ano e mostrar as faixas de nota que cada total de acertos produziu na prática.

6. Mitos e fatos sobre a TRI

Mito: "Quem erra questão fácil é penalizado."

Fato: Não há penalidade automática. O modelo pondera o padrão inteiro: um acerto difícil sem os acertos fáceis esperados vale menos, porque a explicação estatística mais provável passa a ser o acerto casual. Mas nenhuma questão, sozinha, "derruba" a nota.

Mito: "Acertar questão difícil aumenta mais a nota."

Fato: Depende do padrão. Um acerto difícil coerente com o restante do cartão vale muito; o mesmo acerto, incoerente, vale pouco. E questão certa nunca diminui a nota.

Mito: "A TRI torna a prova mais difícil."

Fato: Não. Segundo o próprio guia do INEP, a TRI não altera a dificuldade da prova nem muda significativamente a classificação dos participantes em relação ao percentual de acertos; ela detalha melhor as notas e evita empates em massa. Quem está mais preparado continua se saindo melhor.

Mito: "Minha nota depende do desempenho dos outros."

Fato: Não. A escala usa a média dos concluintes regulares de 2009 apenas como referência histórica. Sua nota depende do seu padrão de respostas e dos parâmetros das questões da sua edição.

Mito: "Dá para prever a nota contando acertos."

Fato: Não existe fórmula de conversão. O que existe são faixas observadas ano a ano - úteis como referência, nunca como promessa.

Mito: "Melhor deixar em branco do que arriscar."

Fato: Ao contrário. Questão em branco é corrigida como errada, e questão certa sempre aumenta a nota - mesmo quando o acerto parece casual ao modelo. A orientação oficial é responder tudo.

Mito: "A redação também é corrigida pela TRI."

Fato: Não. A redação vai de 0 a 1.000 pontos, corrigida por dois avaliadores independentes, com terceira correção quando há discrepância grande entre eles.

7. O que isso muda na sua preparação

- Base em dia vale mais do que "questão difícil decorada". A TRI premia a coerência: construir o edifício do térreo para cima rende mais do que pular andares.
- Nunca deixe questão em branco. Uma questão certa sempre melhora a nota estimada - não por pontos fixos, mas porque melhora o seu padrão de respostas; em branco, ela é corrigida como errada. No fim da prova, marque tudo.
- Desconfie de simuladores e tabelas de conversão. Use faixas observadas como referência, sabendo que a sua nota real vai depender do seu padrão de acertos.
- Treine com provas reais completas, no tempo real. A TRI reage ao padrão global de respostas - e é esse padrão, com fáceis e difíceis misturadas, que você treina num simulado inteiro.

- Não se assuste com a nota crua. Uma nota 550 não é "55% da prova". É uma posição na escala, acima da média de referência de 2009 - e só faz sentido lida assim.

8. Como ler as tabelas de acertos e nota do nosso levantamento

No material irmão deste - o levantamento das 990 questões de Matemática do ENEM de 2015 a 2025 - publicamos, para cada ano e cada total de acertos, a nota mínima, a média e a máxima observadas entre os participantes regulares com nota positiva.

Agora você já sabe ler a dispersão que aparece ali: a distância entre a nota mínima e a máxima de uma mesma linha é a TRI em ação. São alunos com o mesmo número de acertos e padrões de resposta diferentes. A média dá o centro da faixa; os extremos mostram até onde a coerência (ou a falta dela) empurrou a nota naquele ano.

Use essas tabelas como mapa de faixas reais, não como previsão. Elas dizem onde as notas caíram de fato, edição por edição - e não onde a sua vai cair. Essa honestidade é o que separa um levantamento de dados de uma "tabela de conversão" de internet.

Fontes oficiais e referências

INEP - Nota técnica: Procedimento de cálculo das notas do Enem (dezembro de 2011)
Documento oficial que detalha o modelo logístico de três parâmetros, o método EAP, a escala ancorada em 2009 e a tripla conferência independente dos resultados (Cespe/UnB, Cesgranrio e INEP).

INEP - Entenda a sua nota no Enem: Guia do Participante (2012/2013)
Cartilha oficial com a explicação didática da TRI, o exemplo dos dois participantes com os mesmos acertos, a analogia das barreiras e as fórmulas do modelo.

INEP - Entenda a Teoria de Resposta ao Item (TRI), utilizada no Enem (nota técnica, 2012)
Apresenta a motivação da TRI, a comparação com a Teoria Clássica dos Testes e o papel dos três parâmetros e do pré-teste.

INEP - Método de correção das provas reconhece o conhecimento e a coerência do participante (notícia, 2019)
Explica por que a anulação de itens ou provas com itens diferentes não comprometem a estimação das notas.

MEC - Conheça todas as etapas para concepção e aplicação das provas do Enem (2026)
Visão atual do exame, incluindo o papel da TRI e da coerência do padrão de respostas na correção.

Microdados do ENEM (INEP, edições 2015 a 2025)
Base oficial de resultados usada no levantamento irmão deste material; dela sai, por exemplo, a nota máxima de 1.008,3 em Matemática no ENEM 2015.

Literatura técnica clássica
ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. Teoria de Resposta ao Item: conceitos e aplicações (2000). BAKER, F. B.; KIM, S. Item Response Theory (2004). BIRNBAUM, A. Some latent trait models... (1968).

Material gratuito do Politano Aprova. Companheiro do levantamento "O que mais cai em Matemática no ENEM (2015-2025)". Texto conferido contra os documentos oficiais listados acima; os exemplos numéricos são ilustrativos e não reproduzem o cálculo oficial.