

LETRAMENTO ESTATÍSTICO: O QUE SABEM OS ESTUDANTES BRASILEIROS DE ECONOMIA SOBRE ANÁLISE DE REGRESSÃO

Statistical Literacy: what brazilian economics students know about regression analysis

Diêgo Bezerra de Melo Maciel

Gilda Lisbôa Guimarães

Resumo

Nos cursos de Economia o conteúdo da Análise de Regressão possui grande relevância, sendo, inclusive, tratado em uma disciplina específica denominada Econometria. No entanto, desde os anos de 1980, são relatadas, internacionalmente, dificuldades de aprendizagem nesse conteúdo. Esta pesquisa buscou analisar o que sabem os estudantes brasileiros de Economia sobre Análise de Regressão, sob a perspectiva do modelo de Letramento Estatístico. Para isso, aqui foi analisada uma questão apresentada no principal instrumento nacional de acompanhamento da aprendizagem superior: o ENADE. Constatou-se a exigência de múltiplas habilidades, as quais foram ao encontro do modelo de Letramento Estatístico adotado. Porém, observou-se que menos da metade dos participantes (43,75%) obtiveram êxito na questão. Isso mostra a necessidade urgente de mudanças no processo de ensino, dotando o economista brasileiro com as habilidades estatísticas necessárias para o trabalho empírico com os dados, uma de suas principais demandas profissionais.

Palavras-chave: Educação Estatística; ENADE; Ensino Superior.

Abstract

In Economics courses, the content of Regression Analysis has great relevance, being even treated in a specific subject called Econometrics. However, since the 1980s, learning difficulties in this content have been reported internationally. This research sought to analyze what Brazilian Economics students know about Regression Analysis, from the perspective of the Statistical Literacy model proposed by Gal (2002). For this, here an issue presented in the main national instrument for monitoring higher learning was analyzed: ENADE. There was a requirement for multiple skills, which were in line with the adopted Statistical Literacy model. However, it was observed that less than half of the

participants (43.75%) were successful in the question. This shows the urgent need for changes in the teaching process, providing Brazilian economists with the necessary statistical skills for empirical work with data, one of their main professional demands.

Keywords: Statistical Education; ENADE; Higher education.

Introdução

A trajetória de desenvolvimento da Ciência Econômica tem sofrido profundas mudanças desde a oficialização de seu surgimento, em 1776, com os trabalhos do filósofo britânico Adam Smith. As rebuscadas teorias estereotipadas do século XVIII, as quais tentavam explicar os determinantes do crescimento econômico dos países, cederam espaço, a partir de meados do século XIX, para modelos matemáticos e/ou estatísticos, os quais possuíam (possuem) a difícil missão metodológica de modelar, explicar e/ou prever os mais variados fenômenos econômicos.

A principal repercussão desse movimento para a formação inicial do economista é a exigência de uma sólida formação quantitativa. No caso do Brasil, de acordo com as Diretrizes Curriculares Nacionais do Curso de Graduação em Economia (BRASIL, 2007), os cursos devem possibilitar, dentre outras competências e habilidades, a utilização de “*formulações matemáticas e estatísticas na análise dos fenômenos socioeconômicos*”. Além disso, conforme o mesmo documento, essas graduações devem assegurar, em seus projetos pedagógicos, conteúdos mais

avançados da matemática, da estatística e da econometria.

A Estatística Moderna¹ está presente nos currículos de grande parte dos cursos do Ensino Superior ao redor do mundo, com diferentes níveis de complexidade. A maior inserção científica da Estatística nas diversas ciências justifica parte desse cenário, o qual inicia-se com maior vigor a partir da primeira metade do século XX, com a propagação das contribuições de Ronald Fisher.

Esse quadro originou preocupações institucionais com o ensino da Estatística, notadamente nos Estados Unidos. Assim, a ASA (*American Statistics Association*) – fundada em 1839 como entidade para suporte estatístico governamental, criou, em 1948, a comissão de Educação do ISI (*International Statistical Institute*). O ISI passou a organizar conferências internacionais sobre o ensino da Estatística, denominadas ICOTS (*International Conference on Teaching Statistics*), realizadas a cada quatro anos desde 1982² e em vigor até o presente momento. Em 1991, foi criada, como uma seção de Educação do ISI, a IASE (*International Association for Statistical Education*), objetivando promover a melhoria da Educação Estatística em escala mundial. A IASE é reconhecida, atualmente, como uma organização internacional da Educação Estatística.

Os estudantes dos cursos de Economia estão imersos em currículos que apresentam, em geral, a maior quantidade de disciplinas e conteúdos relacionados com a Estatística, quando comparados aos demais cursos superiores, excetuando-se apenas o bacharelado em Estatística. Dentro desse contexto, é indiscutível a grande presença do conteúdo da Análise de Regressão nos currículos dos cursos de Economia.

Entretanto, apesar do protagonismo da Estatística nesses cursos, são muitas as dificuldades de aprendizagem apresentadas pelos estudantes. Em específico, citam-se as barreiras de compreensão acerca dos

conceitos estatísticos relacionados com a Análise de Regressão, as quais já vêm sendo relatadas, internacionalmente, há muito tempo (SOWEY, 1983; KENNEDY, 1998a; KENNEDY, 1998b; BECKER e GREENE, 2001; CRAFT, 2003; BEKKERMAN, 2015; ANGRIST e PISCHKE, 2017; KASSEN, 2019; ARKES, 2020).

Em um estudo recente com estudantes chineses de Economia, Agyeman et al. (2021) constaram que os alunos reconhecem a importância da Estatística para sua formação. Entretanto, a maior parte deles não se sente segura, quanto aos conhecimentos matemáticos e computacionais exigidos, além da falta de compreensão dos conceitos ligados à própria Economia. Além disso, os autores observaram elevados níveis de ansiedade e falta de motivação entre os graduandos de Economia.

Em termos nacionais, são poucos os estudos debruçados sobre as ações de aprendizagem desenvolvidas nos cursos de Economia, notadamente com relação à Análise de Regressão. Tem-se conhecimento, até então, das investigações de Pagliarussi (2018), que propõe um conjunto de simulações de Monte Carlo para facilitar a aprendizagem sobre distribuição amostral, no contexto do modelo clássico de Regressão Linear (Simple e Múltipla).

Pelo exposto, destaca-se a importância de investigar o que sabem sobre conceitos estatísticos estudantes brasileiros de Economia, especialmente no que diz respeito ao conteúdo da Análise de Regressão. Para este trabalho, estudar esses indicadores significa, principalmente, compreender o conhecimento estatístico conceitual apresentado por esses estudantes, ou seja, prioriza-se, aqui, a aprendizagem voltada para a “*compreensão dos princípios e relações gerais, (...) compreensão do porquê das estatísticas, além do como*” (CROOKS; BARTEL e ALIBAL, 2019, p. 46).

Esse cenário de aprendizagem possui íntima relação com a perspectiva do

¹ Estatística Moderna é um termo utilizado para designar a sistematização da Estatística, na forma como se conhece nos dias atuais. Esse processo inicia-se na primeira metade do século XX e reflete, principalmente, as importantes contribuições de Ronald Fisher. Assim, em termos curriculares, a Estatística Moderna insere-se sob a forma de três grandes blocos de conteúdo, nem sempre tratados

conjuntamente nos cursos superiores: Análise Descritiva, Probabilidade e Inferência Estatística.

² A partir de 1986, a ICOTS dedica uma seção voltada para discussão da Educação Estatística na área de Economia e Negócios. Em algumas edições, contemplaram-se trabalhos específicos na área de Econometria.

Letramento Estatístico (LE) descrito por Gal (2002), o qual, em linhas gerais, valoriza a necessidade e capacidade do sujeito em interpretar criticamente a informação estatística, não apenas pela valorização com fórmulas e algoritmos matemáticos. Nesse sentido, o referido autor sistematizou um conjunto de conhecimentos/habilidades necessárias para um adulto (com educação básica concluída) ser letrado estatisticamente.

Diante disso, para investigar os tipos de habilidades desenvolvidas na aprendizagem estatística dos estudantes de Economia, optamos pela análise das questões do Exame Nacional de Desempenho dos Estudantes (ENADE). Instituído a partir de 2004, o ENADE é o principal instrumento nacional de avaliação e acompanhamento da aprendizagem superior. O Exame avalia os estudantes, segundo os conteúdos, competências e habilidades previstos nas diretrizes curriculares dos cursos. Faz parte do componente curricular obrigatório, sendo aplicado, periodicamente, em ciclos de três anos, para estudantes concluintes.

No caso dos cursos de Economia, a última edição do ENADE ocorreu em 25 de novembro de 2018. O Exame foi destinado para 9.580 alunos com expectativa de conclusão até julho de 2019, os quais estavam distribuídos entre 195 cursos de graduação em Economia de todo o país. Nesse Exame, 30% das questões objetivas do componente específico do curso (8, em termos absolutos) relacionaram-se, exclusivamente, com conteúdos estatísticos. Desse total, 50% (quatro questões) correspondeu à Análise de Regressão, reforçando a importância do referido conteúdo na formação estatística do futuro economista.

Assim, essa pesquisa analisa uma das questões do ENADE/Economia, realizado em 2018, sobre Análise de Regressão. Essa análise é realizada para cada um dos distratores da questão, sob o ponto de vista dos diferentes tipos de conhecimento/habilidades estatísticas elencadas por Gal (2002). Além disso, são mensurados os níveis de desempenho apresentados pelos estudantes na questão.

Por fim, ressalta-se que este estudo é parte de uma pesquisa maior, a qual está sendo desenvolvida para uma Tese de Doutorado, no âmbito do Grupo de Estudos em Educação Estatística (GREF)³ vinculado ao Programa de Pós-graduação em Educação Matemática e Tecnológica da Universidade Federal de Pernambuco.

Letramento Estatístico

As perspectivas de ensinar e aprender estatística vêm sofrendo importantes mudanças, notadamente, a partir dos anos 2000, com a consolidação do campo de pesquisa da Educação Estatística. A partir daí, novos olhares são lançados sobre esse processo. A memorização de fórmulas e algoritmos cede espaço para uma aprendizagem conceitual, tornando o estudante apto a entender, interpretar, criticar e reagir à informação estatística. É o Letramento Estatístico (LE).

Isso posto, em um contexto mais amplo, Gal (2002) propõe um modelo de LE que busca identificar as competências e habilidades necessárias para que um cidadão adulto se torne letrado estatisticamente. Para o autor, o avanço na produção e disseminação de dados torna o LE uma habilidade fundamental para o cidadão, esperando-se que os estudantes ingressem no ensino superior com essa habilidade.

A importância do LE pode ser visualizada sob diversos aspectos cotidianos da vida de um cidadão adulto, os quais estão relacionados com algum tipo de informação estatística. Gal (2002) destaca alguns desses aspectos, no contexto de diversos fenômenos sociais que afetam a sociedade, tais como: taxas de criminalidade, crescimento populacional, disseminação de doenças, ou até mesmo tendências de emprego. O autor avança, afirmando que o LE pode contribuir para melhorar a capacidade do adulto em fazer escolhas, nas situações baseadas no acaso (por exemplo, compra de bilhetes de loteria ou apólices de seguro).

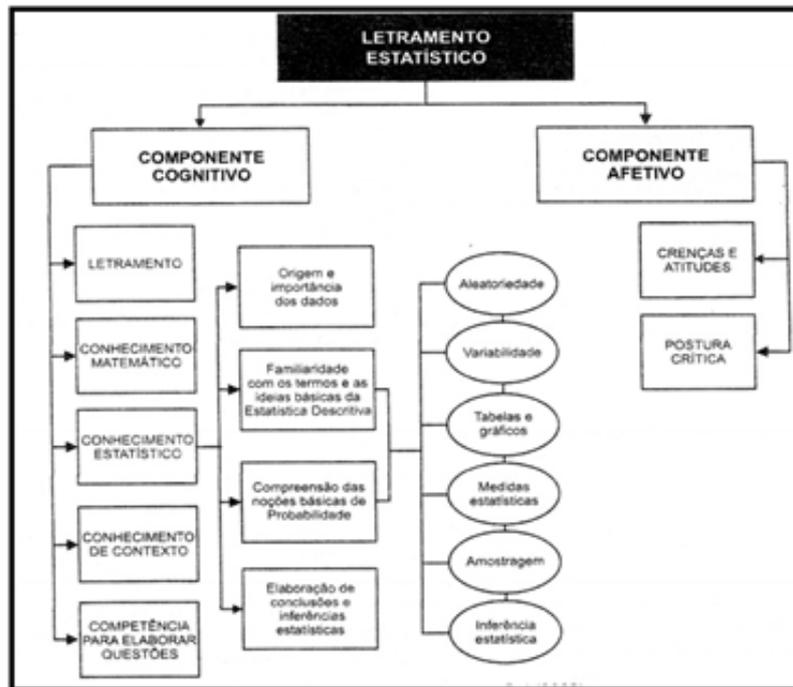
Diante disso, Gal (2002) apresenta um modelo que sistematiza o LE (Figura 1), direcionando-o para as habilidades e posturas que um adulto já escolarizado deva possuir, perante a informação estatística. Em

³ <https://ufpepesquisas.wixsite.com/gref>

verdade, o modelo está calcado sob dois componentes interrelacionados, nomeados como: i) componente cognitivo e ii) componente afetivo. O primeiro objetiva, essencialmente, a articulação entre a capacidade de compreender, interpretar e avaliar criticamente as informações estatísticas. Já o segundo, evidencia como aspectos subjetivos inerentes ao adulto letrado relacionam-se com sua postura diante da informação estatística.

Isso posto, segundo o componente cognitivo, o sujeito letrado precisará ativar, conjuntamente, cinco elementos de conhecimento, a saber: i) letramento; ii) conhecimento matemático; iii) conhecimento estatístico; iv) conhecimento de contexto; e v) competência para elaborar questões (questionamento crítico).

Figura 1: Modelo de letramento estatístico



Fonte: Gal (2002), apud Cazorla e Utsumi (2010, p.12)

Avançando nesses elementos, o “letramento” consiste, grosso modo, em entender a informação estatística, transmitindo-a de forma clara e objetiva, quando necessário. Esse entendimento é feito, na maior parte das vezes, a partir de uma interpretação gráfica e/ou tabular e textual. Já o “conhecimento matemático” está relacionado, principalmente, com a contribuição teórica desse tipo de conhecimento para construção dos conceitos estatísticos. Todavia, o autor argumenta que um número ou expressão pode apresentar sentidos diferentes na matemática e na estatística.

Por exemplo, na Análise de Regressão é fundamental compreender a diferença entre modelo matemático e modelo estatístico. No primeiro caso, funções como

$y=4x+2$ exaurem completamente a relação entre as variáveis. Já no modelo estatístico, a impossibilidade de acesso aos dados populacionais, dentre outras limitações, impede um relacionamento exato entre as variáveis. Por conta disso, em termos estatísticos, a relação, ora apresentada, entre “x” e “y” deve ser posta na forma $Y=4x + 2 + e$, em que “e” – chamado de erro aleatório - reúne todas as informações que não puderam ser enquadradas no modelo sobre Y, agora aleatório. Em outras palavras, o erro aleatório é fruto da natureza amostral do modelo estatístico, fato que impõe um nível de variabilidade para as variáveis observadas.

Ainda no campo cognitivo, Gal (2002) aponta os conhecimentos estatísticos necessários ao LE. Como percebe-se, a partir

da Figura 1, esses conhecimentos abarcam, praticamente, todos os pilares do conteúdo estatístico presentes nos cursos de Economia e demais cursos de nível superior. Porém, Gal (2002) indica que a lista desses conhecimentos estatísticos não é exaustiva, tão pouco universal. Dessa forma, esses conhecimentos “não podem ser discutidos em termos absolutos, pois dependem do nível desejado de letramento estatístico (...) e dos contextos culturais” (GAL, 2002, p. 7).

Por conta disso, o conhecimento estatístico em tópicos avançados da Análise de Regressão pode ser mais apropriado para um economista, enquanto dispensável para a maioria dos cidadãos comuns. Por isso, a extensão e nível de complexidade do conhecimento estatístico dependerá das motivações e objetivos pessoais do adulto perante a informação.

Para finalizar o componente cognitivo, o teórico incorporou os elementos “conhecimento de contexto” e “questionamentos críticos”. O primeiro está relacionado ao contexto no qual a informação estatística está inserida. Assim, de acordo com Gal (2002), a interpretação adequada da mensagem por adultos depende de sua capacidade em colocar mensagens em um contexto, conduzindo à reflexão crítica dos dados. Já o segundo faz referência aos questionamentos acerca da informação recebida, sob diversos aspectos, tais como: confiabilidade da fonte dos dados; forma de coleta da amostra, bem como sua representatividade e período da pesquisa, entre outros.

Por outro lado, o componente afetivo do modelo impõe ao adulto letrado uma postura ativa, eximindo-o da posição de mero consumidor da informação. Esse componente é utilizado pelo autor para agregar três conceitos subjetivamente relacionados, porém, distintos: i) crenças (opiniões individuais); ii) atitudes (sentimento em relação a objetos, ações ou temas); e iii) postura crítica (propensão ao comportamento questionador). Existe, nesse sentido, uma dificuldade teórica e prática em separar esses conceitos, dado o alto grau de imbricamento entre eles.

Ressalta-se, nesse componente, como o comportamento do sujeito, ao nível de suas disposições pessoais, afeta sua postura perante a informação estatística, revelando,

com isso, a formação dos comportamentos de letramento estatístico socialmente convencionados.

Diante da importância do LE, a partir do modelo proposto por Gal (2002), o desafio posto foi analisar o que estudantes de Economia demonstram saber sobre Análise de Regressão. Assim, a próxima seção se propõe a trazer um panorama geral desse processo, com ênfase no cenário internacional, devido à carência de estudos nacionais, conforme relato anterior.

Análise de Regressão nos cursos de Economia: como vão as coisas?

Nos cursos de Economia, o conteúdo da Análise de Regressão possui grande relevância, sendo, inclusive, tratado em uma disciplina específica chamada de Econometria. Tipicamente, nesses cursos, a referida disciplina é ofertada no último ano de formação dos estudantes. Nessa altura, esses acadêmicos já obtiveram contato prévio com conteúdos da Análise Descritiva, Probabilidade e Inferência Estatística.

Todavia, ao analisar os cursos canadenses de Economia, Kennedy (1998a) alerta que, contrariamente ao pensamento dos professores de Econometria, os estudantes não compreendem a lógica estatística por detrás dos Modelos de Regressão. No caso dos cursos estadunidenses, Kassens (2019) relata a mesma percepção de Kennedy (1998b). Assim, a autora destaca que o processo de aprendizagem desenvolvido direciona os futuros economistas para a abstração matemática, em detrimento da compreensão dos principais conceitos estatísticos relacionados com o conteúdo.

Ainda de acordo com a autora, após formados, os acadêmicos de Economia entram no mercado de trabalho com dificuldades de entender e resolver os diversos problemas práticos demandados pelas empresas. Esse quadro é consequência, dentre outros fatores, da “falta de equilíbrio, ou assimetria, na educação econométrica, entre teoria e prática” (KASSENS, 2019, p. 2).

Recentemente, Cladera (2021) relatou, a partir da percepção dos docentes, que os estudantes espanhóis de Economia não possuem atitudes positivas em relação

ao conteúdo da Análise de Regressão. A autora conclui que “*isso pode ser um problema para aprender e adquirir habilidades econométricas*” (CLADERA, 2021, p. 1002).

No tocante aos cursos brasileiros de Economia, o cenário parece não divergir. Assim, Pagliarussi (2018) constata que os alunos não compreendem os conceitos básicos da inferência estatística. Segundo o autor, caso um professor de Econometria aborde a distribuição amostral dos estimadores dos modelos de Regressão, “irá perceber nos olhos dos seus alunos a dificuldade de compreender o seu significado” (PAGLIARUSSI, 2018, p.2). O mesmo autor afirma, sob um alto nível de probabilidade, que um “bom” professor brasileiro de econometria desenvolve “aulas que requerem um nível razoavelmente elevado de raciocínio teórico e matemático” (PAGLIARUSSI, 2018, p.2).

Finalmente, em uma perspectiva mais ampla, Oliveira Júnior; Barros Neto e Alves (2020) reconhecem problemas na aprendizagem da Correlação e da Regressão nos cursos da área de negócios, a partir da análise dos dez livros mais utilizados para ensino desses conteúdos. Os autores concluem a existência de um *modus operandi* comum para apresentação dos conteúdos da Regressão Linear e que a maioria dos manuais não avança nos conceitos estatísticos inferenciais, associados com as estimativas geradas pelos modelos de Regressão.

Uma possível justificativa para todas essas constatações apresentadas pode guardar relação com a diversidade de conceitos estatísticos envolvidos no estudo da Regressão, tornando a aprendizagem desse conteúdo algo não muito simples (BATANERO et al., 2017).

Método

Neste artigo, foi escolhida a Questão 25 presente na última edição do Enade (2018), a qual foi aplicada para 9.580 estudantes do último ano do curso de Economia. Desse total, foram excluídos 1.816 possuidores de nota zero em todas as questões do referido Exame, restando para análise 7.764 indivíduos. Ressalta-se ainda a necessidade de excluir as respostas sem

validade (ilegíveis, em duplicidade, com cor não permitida, etc.). Por isso, no caso da Questão 25, foram descartadas mais 36 respostas (0,46%), acarretando um quantitativo final de 7.728 estudantes analisados.

As questões do ENADE são tratadas aqui, metodologicamente, como situações profissionais vivenciadas pelo Economista, as quais são tomadas sob a perspectiva de um economista (adulto) letrado estatisticamente. As questões (situações) são desenhadas de tal forma que o “conhecimento estatístico” represente o eixo central de suas respectivas análises.

Prosseguindo, o fato desse economista letrado estar circunscrito a situações exclusivamente profissionais, não o exime de possuir juízo de valor sobre a informação veiculada (a própria questão do ENADE). Isso permite a ele (adulto economista) confrontar, por exemplo, a fonte dos dados, a qualidade da informação, possíveis tendenciosidades, erros, omissões etc. É o “questionamento crítico”, nos termos colocados pelo modelo de Gal (2002).

Consequentemente, em um sentido mais amplo, serão aproximados possíveis traços comportamentais esperados desse adulto, perante a informação estatística. Esse comportamento é ancorado pela análise da narrativa dos distratores da questão. Dito de outra forma, certos distratores, ou o próprio enunciado da questão, servirão para representar uma expectativa de postura, a qual será referenciada pelos elementos afetivos do modelo. Esses elementos serão tratados em conjunto, dada a dificuldade de separá-los, conforme já relatado.

Apresentadas as principais estratégias operacionais para identificar as habilidades exigidas pela questão, segundo o modelo de Gal (2002), o caminho metodológico da pesquisa seguiu os seguintes passos: 1) apresentação da questão; 2) análise da solução, por distrator; 3) percentual de escolha dos estudantes em cada distrator e 4) identificação dos tipos de habilidades de LE exigidos, segundo a estratégia operacional aplicada.

Resultados

A Questão 25 do ENADE (Figura 2) abordou a utilização de um modelo de Regressão Linear Simples para investigar a relação entre o nível de desemprego e o crescimento econômico dos países, a partir dos pressupostos teóricos da Lei de Okun. Assim, foi construído um modelo para prever a taxa de desemprego de um país “X” hipotético, com base na sua taxa natural de desemprego e na taxa de crescimento de seu produto (interno bruto). Em seguida, exibiu-se uma tabela, contendo os valores observados para as taxas de desemprego e do

produto, no período de 2013 a 2017, referentes àquele país.

Isso posto, a resolução da questão exigiu, em termos conceituais, a diferenciação entre “valor observado” e “valor previsto”. Esses conceitos são fruto do erro aleatório relacionados com as variáveis dependentes. A partir disso, o valor previsto (da taxa de desemprego) de cada ano é obtido a partir da aplicação de operações matemáticas básicas, por meio dos valores das taxas de crescimento do produto aplicadas ao modelo.

Figura 2 - Questão 25 – ENADE, 2018

QUESTÃO 25

Um pesquisador resolveu estimar uma versão da Lei de Okun para determinado país X. O resultado é apresentado na equação a seguir.

$$u_t = u_n - 0,5 gy_t + c_t;$$

em que u_t é a taxa de desemprego observada para o ano t ; u_n é a taxa de desemprego natural; gy_t é a taxa de crescimento do produto no ano t ; c_t é o termo de resíduo. O país apresenta uma taxa de desemprego natural igual a 10%.

Com o objetivo de analisar a predição do modelo, esse pesquisador utilizou os dados a seguir, para alguns anos selecionados.

Dados anuais selecionados do país X

Ano	Taxa de Crescimento do Produto	Taxa de Desemprego Observada
2013	4%	6%
2014	8%	5%
2015	4%	7%
2016	2%	10%
2017	10%	5%

Considerando as informações apresentadas, assinale a opção correta.

- A** Para o ano de 2013, o modelo previu uma taxa de desemprego inferior à observada.
- B** Para o ano de 2014, a taxa de desemprego estimada foi igual à observada.
- C** Para o ano de 2015, o modelo superestimou a taxa de desemprego.
- D** Para o ano de 2016, o erro de previsão do modelo foi igual a zero.
- E** Para o ano de 2017, o erro de previsão do modelo foi positivo.

Fonte: Instituto de Pesquisas Educacionais Anísio Teixeira – INEP.

Analisando os distratores da questão, percebe-se, em todos eles, um direcionamento para o confronto entre as taxas de desemprego previstas pelo modelo com aquelas observadas (“reais”). Desta feita, foram abordados aspectos atrelados à qualidade preditiva do modelo para as taxas de desemprego do país hipotético, no

período de 2013 – 2017, utilizando as taxas de crescimento do produto (gy_t) como variável explicativa.

A partir disso, o distrator “A” informou o seguinte: “para o ano de 2013, o modelo previu uma taxa de desemprego inferior à observada”. O distrator está falso. Segundo o modelo, a taxa de desemprego

estimada para o referido ano seria de 8% ($10\% - 0,5 \times 4\%$). Já o valor observado dessa taxa – no ano de 2013, foi de 6%. Com isso, o modelo estimou uma taxa maior que a observada, e não menor.

Já o distrator “B” confrontou os dados estimados pelo modelo com os observados para o ano de 2014. Assim, afirmou que “a taxa de desemprego estimada foi igual à observada”. O distrator está falso. Nota-se, para 2014, uma taxa de desemprego estimada em 6% ($10\% - 0,5 \times 8\%$), enquanto observou-se 5%. Nesse caso, as respectivas taxas foram diferentes.

Em relação ao ano de 2015, o distrator “C” declarou que “o modelo superestimou a taxa de desemprego”. O distrator é verdadeiro. Constatou-se para o referido ano uma estimativa de 8%, ($10\% - 0,5 \times 4\%$), a despeito da taxa observada (7%). Isso posto, houve uma superestimação na ordem de 1% ($8\% - 7\%$), tal como foi afirmado.

O distrator “D” informou não ter havido erro de previsão da taxa de desemprego para o ano de 2016 (“Para o ano de 2016, o erro de previsão do modelo foi igual à zero.”) O distrator está falso. Percebeu-se uma diferença na ordem de 1% entre a taxa estimada (9%) e a observada (10%).

Finalmente, o distrator “E” informou que, “para o ano de 2017, o erro de previsão do modelo foi positivo”. O distrator está falso. Nesse ano específico, o erro de previsão do modelo foi zero, ou seja, a taxa estimada ($10\% - 0,5 \times 10\% = 5\%$) foi igual à observada (5%).

O Quadro 1 apresenta as escolhas dos estudantes para cada distrator da questão. Nota-se um desempenho geral baixo, pois menos da metade dos graduandos (43,75%) optou pelo distrator correto “C”. Isso indica

que boa parte desses estudantes possuem dificuldades com conceitos básicos referentes ao conteúdo da Regressão Linear. Essas dificuldades podem ser reflexo do ensino praticado nos cursos brasileiros de Economia, o qual parece optar por tópicos mais avançados do conteúdo, em detrimento das definições conceituais mais básicas.

Um outro possível fator reside na ausência de uma perspectiva prática, em sala de aula, para avaliar a qualidade de um modelo. Tipicamente, os manuais utilizados suportam a validação de um modelo apenas sob o cálculo de testes de hipóteses, obtenção de coeficientes de determinação, erros-padrão.

Dessa forma, do ponto de vista pedagógico, os erros de estimativa figuram apenas em dimensão teórica, sem o trabalho empírico de validação com os dados observados. Sobressalta-se, tão somente, o aspecto estatístico desses erros; no jargão: “possuem valor esperado zero e variância constante”.

Assim, conceitos como superestimação e subestimação podem soar pouco familiares aos estudantes, constituindo uma importante fonte para que não obtenham êxito na questão. Fato reforçado pela semelhança entre os distratores: a abordagem foi a mesma em todos eles, mudando-se, apenas, um ou outro valor ou período de análise. Esses fatores, em conjunto, impõem ao estudante brasileiro uma visão limitada do conteúdo, acarretando dificuldades perante situações práticas, tal como levantado por Kassens (2019), em suas análises relacionadas com a aprendizagem dos estudantes de Economia norte-americanos.

Quadro 1 – Distribuição absoluta e percentual das escolhas dos estudantes, segundo os distratores da questão 25 - ENADE/2018

DISTRATOR	ESTUDANTES
A) Para o ano de 2013, o modelo previu uma taxa de desemprego inferior à observada.	1.525 (19,64%)
B) Para o ano de 2014, a taxa de desemprego estimada foi igual à observada.	722 (9,22%)
C) Para o ano de 2015, o modelo superestimou a taxa de desemprego.	3.397 (43,75%)
D) Para o ano de 2016, o erro de previsão do modelo foi igual à zero.	1.009 (13%)
E) Para o ano de 2017, o erro de previsão do modelo foi positivo.	1070 (13,78%)

Fonte: Elaboração Própria, a partir dos dados do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira - INEP

A partir da análise de cada distrator, detalham-se, a seguir, as habilidades de LE envolvidas nessa situação profissional (questão), segundo a estratégia operacional desenvolvida para este Estudo:

“Letramento”: notou-se a necessidade de conferir sentido econômico a um modelo de regressão linear simples, a partir de sua formulação matemática. Em específico, foi preciso compreender a aplicabilidade desse modelo a um conjunto de dados observados, transmitindo, com isso, opiniões sobre a qualidade das estimativas geradas.

“Conhecimento estatístico”: domínio dos pressupostos teóricos dos modelos de regressão linear, notadamente, os relacionados com os erros de previsão. Nessa esfera, valorizou-se a compreensão acerca das consequências práticas dos erros de previsão sobre a qualidade das estimativas geradas; em específico, os conceitos de superestimação e subestimação. Percebeu-se, ainda, a diferenciação conceitual entre variável dependente e independente, sendo isso a “primeira ideia importante na regressão” (BATANERO et al., 2017).

“Conhecimento matemático”: apesar do conteúdo exigir, *per si*, habilidades com elementos do cálculo diferencial, a situação representada pela questão requereu, tão somente, a utilização

de operações básicas com dados decimais e percentuais. Essas operações foram realizadas para obtenção das estimativas das taxas de desemprego, bem como para as comparações dessas estimativas com as taxas observadas.

“Conhecimento de contexto”: esse elemento esteve calcado sobre uma teoria econômica conhecida por Lei de Okun. Essa teoria - formulada na década de 1960, pelo economista Arthur Okun - afirma haver uma estreita relação negativa entre os níveis de desemprego e a taxa de crescimento econômico (produto) dos países. Assim, em geral, ao longo do tempo, observou-se que níveis menores nas taxas de desemprego dependiam de taxas maiores de crescimento econômico, tudo o mais constante.

“Questionamento Crítico”: a situação desenhada não apresentou nenhuma informação sobre os aspectos inferenciais do modelo estimado. Espera-se de um economista letrado a capacidade de questionar essa ausência, pois isso compromete o embasamento estatístico das opiniões acerca da qualidade preditiva do modelo. Fora isso, é possível contestar a falta do período temporal da amostra utilizada. Esse aspecto é importante, pois, se o recorte de tempo amostrado estiver muito distante do período confrontado (2013-2017), os resultados das previsões ficam seriamente comprometidos.

“Elementos afetivos”: a apresentação de um modelo de regressão sem os devidos resultados inferenciais evidencia a típica crença em tecer conclusões gerais sem o conhecimento sobre a capacidade estatística do modelo. Assim, não é raro acreditar na fiabilidade dos resultados apenas sob o ponto de vista da teoria econômica utilizada; no caso em tela, a Lei de Okun. Conseqüentemente, os aspectos práticos dos erros de estimativa da regressão são postos em segundo plano, considerando-se tais erros apenas em seu sentido teórico.

Conclusões

Essa pesquisa buscou investigar a aprendizagem estatística dos estudantes brasileiros de Economia sobre o conteúdo da Análise de Regressão, utilizando uma questão (Questão 25) da última edição do ENADE para os cursos de Economia. A partir disso, o estudo foi conduzido sob duas frentes: i) interpretação dos tipos de habilidades exigidas pela questão, a partir do modelo de LE proposto por Gal (2002); e ii) mensuração dos níveis de desempenho apresentados pelos estudantes.

Constatou-se, assim, que a referida questão representou um cenário profissional, no qual habilidades com fórmulas e algoritmos estatísticos não eram suficientes. O estudante precisou ir mais além: foi necessário emitir opinião sobre a qualidade preditiva de um determinado modelo, formulado com base em uma Teoria Econômica (Lei de Okun).

Em termos de conhecimento matemático, demandou-se apenas operações matemáticas triviais com porcentagens e números decimais. O conhecimento estatístico restringiu-se aos conceitos de variável dependente/independente e no cálculo/análise dos erros de predição.

Tomando a questão analisada como um simulacro de uma determinada situação profissional, foi possível estabelecer alguns questionamentos críticos, notadamente relacionados com a ausência de testes estatísticos inferenciais. Esse aspecto, certa forma, está associado com a crença na “confiança cega” no modelo econométrico, em detrimento da sua real capacidade de explicar a realidade econômica fática.

Ainda que esses elementos afetivos não estejam vinculados, diretamente, com o desempenho do aluno na questão do ENADE, tomou-se a sua existência sob uma perspectiva mais ampla, cultural. Eles refletem, certa forma, as interações dos estudantes com seus professores de Estatística; estes no papel de produtores e disseminadores de certas “visões de mundo” acerca dos mais diversos tipos de situações e contextos estatísticos.

Assim, com base nesses levantamentos, a questão exigiu uma aprendizagem múltipla, pautada na comunhão entre diferentes tipos de habilidades, as quais transcendem a aplicação do conhecimento técnico matemático. Caminhou-se para uma leitura estatística conceitual do Modelo de Regressão, na perspectiva do Letramento Estatístico, nos moldes interpretativos do modelo preconizado por Gal (2002).

Em paralelo, observou-se que menos da metade dos estudantes avaliados (43,75%) obtiveram êxito na questão. Esse cenário pode oferecer importantes pistas acerca de como vem ocorrendo o processo de aprendizagem estatística dos estudantes brasileiros de Economia. Há indícios de que as fragilidades identificadas no contexto internacional ganham eco na realidade brasileira. Isso mostra a necessidade de adoção urgente de processos de ensino que abordem a perspectiva do Letramento Estatístico, dotando o economista brasileiro com as habilidades necessárias para o trabalho empírico com os dados, uma de suas principais demandas profissionais.

Referências

- AGEYEMAN, F.O.; SAMPENE, A. K.; ZHIQIANG, M.; LI, C.; LI, M.; ROBERT, B.; WIREDU, J. From Theory to Practice of Promoting Student Interest in Econometrics: Application of Structural Equation Model. **International Journal of Development Research**, v.11, n.11, p.52124-34, nov.2021.
- ANGRIST, J. D.; PISCHKE, J-S. Undergraduate Econometrics Instruction: Through Our Classes, Darkly. **Journal of Economic Perspectives**, Pittsburgh, v. 31, n. 2, p. 125-44. 2017.
- ARKES, J. Teaching Graduate (and Undergraduate) Econometrics: Some Sensible

Shifts to Improve Efficiency, Effectiveness, and Usefulness. **Econometrics**, Basel, v. 8, n. 3, p. 36, set. 2020.

BATANERO, C.; GEA, M. M.; LÓPEZ-MARTÍN, M. M.; ARTEAGA, P. Análisis de los conceptos asociados a la correlación y regresión en los textos de bachillerato. **Didacticae**, Barcelona, v. 1, n. 1, p. 60-76. 2017.

BECKER, W. E.; GREENE, W. H. Teaching statistics and econometrics to undergraduates. **The Journal of Economic Perspectives**, Pittsburgh, v. 15, n. 4, p. 169-182. 2001.

BEKKERMAN, A. The role of simulations in econometrics pedagogy. **Wiley Interdisciplinary Reviews: Computational Statistics**, [s.l.] v. 7, n. 2, p. 160-165, mar./abr. 2015.

BRASIL. **Diretrizes Curriculares Nacionais do Curso de Graduação em Economia**. Brasília, MEC, 2007.

CAZORLA, I.; UTSUMI, C. M. Reflexões sobre o ensino da estatística na educação básica. In: CAZORLA, I.; SANTANA, E. (Org.). **Do tratamento da informação ao letramento estatístico**. 1 ed. Itabuna: Via Litterarum, 2010.

CLADERA, M. Assessing the attitudes of economics students towards econometrics. **International Review of Economics Education**, Palma de Mallorca, v. 37, abr. 2021.

CRAFT, R. K. Using spreadsheets to conduct Monte Carlo experiments for teaching introductory econometrics. **Southern Economic Journal**, New York, v. 69, n. 3, p. 716-735, jan.. 2003.

CROOKS, N. M.; BARTEL, A. N.; ALIBAL, M. W. Conceptual Knowledge of Confidence

Intervals in Psychology Undergraduate and Graduate Students. **Statistics Education Research Journal**, California, v. 18, n. 1, p. 46-62, 2019.

GAL, I. Adults' statistical literacy: meanings, components, responsibilities. **International Statistical Review**, Netherlands, v. 70, n.1, p. 1-50, 2002.

KASSENS, A. L. Theory vs. practice: Teaching undergraduate econometrics. **The Journal of Economic Education**, [s.l.], v. 50, n. 4, p. 367-370, out. 2019.

KENNEDY, P. E. Using Monte Carlo studies for teaching econometrics. In: BECKER, W. E.; WATTS, M. (Eds.). **Teaching Economics to Undergraduates: Alternatives to Chalk and Talk**. 1 ed. Northampton: Edward Elgar, 1998a. p. 141-159.

KENNEDY, P. E. Teaching undergraduate econometrics: a suggestion for fundamental change. **American Economic Review**, Pittsburgh, v. 88, n. 2, p. 487-491, 1998b.

OLIVEIRA JÚNIOR, A. P.; BARROS NETO, D. F.; ALVES, G.C.S. Estudo sobre a correlação e a regressão linear em livros didáticos do ensino superior no Brasil. **Educação Matemática em Revista**, Porto Alegre, v.1, n.21, p. 128-40, 2020.

PAGLIARUSSI, M. S. O ensino do modelo clássico de regressão linear por meio de simulação de Monte Carlo. **Revista de Contabilidade e Organizações**, Ribeirão Preto, v. 12:e152100, dez. 2018.

SOWEY, E. R. University teaching of econometrics: A personal view. **Econometric Reviews**, [s.l.], v. 2, n. 2, p. 255-333, 1983.

Diêgo Bezerra de Melo Maciel: Doutorando em Educação Matemática e Tecnológica. Universidade Federal de Pernambuco - Recife, Pernambuco – Brasil. diego.mmaciell@ufpe.br Orcid: <https://orcid.org/0000-0002-1198-854X>

Gilda Lisbôa Guimarães: Doutora em Psicologia Cognitiva. Universidade Federal de Pernambuco - Recife, Pernambuco – Brasil. gilda.lguimaraes@gmail.com Orcid: <https://orcid.org/0000-0002-1463-1626>