نامه به سردبیر

معرفی روش محقق‌ساخته جهت ارزیابی کیفیت ارزشیابی دانشجویان از طریق نمرات ثبت شده

مجید رحیمی، فرزانه محمدی*، حسین شهنازی، مهناز شاکریان، مریم نصیریان

مجله ایرانی آموزش در علوم پزشکی/ آذر ۱۴۰۴؛ ۲۵ (۷): ۶۵ تا ۶۷

ارزشیابی دانشجویان یکی از مهم‌ترین مراحل در فرآیند آموزش دانشگاهی است که نقش تعیین‌کننده‌ای در سنجش میزان تحقق اهداف آموزشی و یادگیری دارد. در طول یک نیم‌سال تحصیلی، استاد با بهره‌گیری از روش‌های متنوع آموزشی، مفاهیم نظری و عملی را به دانشجویان منتقل می‌کند و در پایان، با استفاده از ابزارهای مختلف ارزشیابی، میزان یادگیری آنان را مورد سنجش قرار می‌دهد(۱). این ابزارها می‌تواند شامل آزمون‌های میان‌ترم و پایان‌ترم، آزمون‌های عملی و نظری، ارائه سمینار، انجام پروژه‌های تحقیقاتی یا مشارکت در فعالیت‌های کلاسی باشد. برآیند تمامی این ارزشیابی‌ها در نهایت به‌صورت یک نمره نهایی ثبت می‌شود که به‌عنوان نماینده عملکرد دانشجو در آن درس تلقی می‌گردد(۲).

با وجود اهمیت نمره نهایی در تصمیم‌گیری‌های آموزشی و پژوهشی، کم‌تر به کیفیت آن به‌عنوان خروجی ارزشیابی توجه شده است. نمره نهایی، به‌ویژه در آزمون‌های تشریحی یا عملی، ممکن است تحت تأثیر قضاوت فردی استاد، سوگیری یا عدم یکنواختی در تصحیح قرار گیرد(۳). این چالش‌ها ضرورت طراحی و استفاده از روش‌هایی را مطرح می‌کند که بتواند کیفیت ارزشیابی انجام‌شده را از طریق تحلیل نمرات ثبت‌شده بررسی و ارزیابی کند.

نویسندگان،روشی محقق‌ساخته معرفی می‌نمایند که با بهره‌گیری از معیارهای آماری و تحلیلی، کیفیت ارزشیابی استاد را بررسی می‌کند. این روش تحلیل آزمون در انواع دوره‌های آموزشی قابل اجرا بوده و با فراهم‌سازی، امکان مقایسه‌پذیری، به بهبود انسجام و اعتبار فرایند ارزشیابی کمک می‌نماید. بر اساس مطالعات انجام شده چهار معیار جهت ارزیابی کیفیت ارزشیابی دانشجویان از طریق نمرات ثبت شده تعیین شد که عبارتند از:

۱- ضریب تمیز ارزشیابی بر اساس نمرات ثبت شده

۲- ضریب دشواری ارزشیابی بر اساس نمرات ثبت شده

۳- بررسی توزیع آماری نمرات ثبت شده

۴- معنادار بودن اختلاف میانگین در دو گروه نمرات کم‌تر و بیش‌تر از میانگین در نمرات ثبت شده

در نهایت به هر معیار امتیازی بین ۰ تا ۱ تعلق گرفته و هر آزمون در مجموع امتیازی بین ۰ تا ۴ کسب خواهد کرد. برای

* **نویسنده مسؤول:** دکتر فرزانه محمدی (استادیار)، گروه مهندسی بهداشت محیط، دانشکده بهداشت، دانشگاه علوم پزشکی اصفهان، ایران. fm_1363@hlth.mui.ac.ir دکتر مجید رحیمی (دانشیار)، گروه آموزش بهداشت و ارتقاء سلامت، دانشکده بهداشت، دانشگاه علوم پزشکی اصفهان، ایران. (mostafavi@hlth.mui.ac.ir)؛ دکتر حسین شهنازی (استاد)، گروه آموزش بهداشت و ارتقاء سلامت، دانشکده بهداشت، دانشگاه علوم پزشکی اصفهان، ایران. (h-shahnazi@hlth.mui.ac.ir)؛ دکتر مهناز شاکریان (استادیار)، گروه مهندسی بهداشت حرفه‌ای، دانشکده بهداشت، دانشگاه علوم پزشکی اصفهان، ایران. (shakerian.mahnaz@gmail.com)؛ دکتر مریم نصیریان (دانشیار)، گروه آمار و اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی اصفهان، ایران. (maryamnasirian17@gmail.com)

تاریخ دریافت مقاله:۱۴۰۴/۸/۶، تاریخ پذیرش: ۱۴۰۴/۹/۲۶

[Downloaded from ijme.mui.ac.ir on 2026-01-30]

استفاده از این روش می‌بایست حداقل ۱۵ دانشجو در دوره حضور داشته باشند. معیارهای انتخاب شده در ادامه معرفی خواهد شد.

- معرفی ضریب تمیز ارزشیابی بر اساس نمرات

این ضریب نشان‌دهنده قدرت آزمون در تمیز بین دانشجویان با عملکرد مختلف است. ضریب تمیز به توانایی آزمون در تعیین تفاوت بین گروه‌های قوی و ضعیف گفته می‌شود. برای محاسبه این ضریب بر اساس نمره، نمرات دانشجویان بعد از مرتب‌سازی به دو گروه تقسیم می‌شود: گروه نمرات بالا (به‌عنوان مثال ۲۵ درصد بالای نمرات) و گروه نمرات پایین (۲۵ درصد پایین نمرات). ضریب تمیز می‌تواند عددی بین ۰ تا ۱۰۰ باشد. هر چه ضریب تمیز سؤال به ۱۰۰ نزدیک‌تر باشد آن سؤال در تمایز دادن بین داوطلبان خوب از بد، بهتر عمل کرده است. برای تعیین ضریب تمیز، آزمون میانگین نمرات ۲۵٪ بالا از میانگین نمرات ۲۵٪ پایین کسر شده و بر دامنه کل نمرات تقسیم خواهد شد. در این بررسی ضریب کم‌تر از ۲۰٪ بدون امتیاز، ضریب بین ۲۰٪ تا ۴۰٪، ۰/۵ امتیاز و ضریب بالاتر از ۴۰٪، ۱ امتیاز خواهد داشت(۴).

- ضریب دشواری ارزشیابی بر اساس نمرات ثبت شده (Exam Difficulty Index)

از محاسبه ضریب دشواری ارزشیابی می‌توان نتیجه گرفت که چه تعدادی از داوطلبان توانسته‌اند به سوالات پاسخ صحیح بدهند، اگر همه داوطلبان به سوالات پاسخ صحیح بدهند آن آزمون بسیار آسان بوده و اگر هیچ یک از داوطلبان پاسخ ندهند آن آزمون بسیار دشوار است. ضریب دشواری عددی بین صفر تا ۱۰۰ است. ضریب دشواری ارزشیابی در این روش از تقسیم میانگین نمرات حاصل شده به نمره کل ارزشیابی محاسبه می‌شود. برای امتیازدهی به ضریب دشواری آزمون بین ۰ تا ۱ می‌توان از مرزهای زیر استفاده کرد(۴).

توضیحات	محدوده ضریب دشواری
آزمون بسیار دشوار (امتیاز حاصله =۰)	٪۰-٪۲۰
آزمون دشوار (امتیاز حاصله = ۰/۵)	٪۲۱-٪۳۰
آزمون متوسط (امتیاز حاصله =۱)	٪۳۱-٪۷۰
آزمون آسان (امتیاز حاصله = ۰/۵)	٪۷۱-٪۸۰
آزمون بسیار آسان (امتیاز حاصله =۰)	٪۸۱-٪۱۰۰

- **بررسی توزیع آماری نمرات ارزشیابی**

بررسی توزیع نرمال نمرات در آزمون‌های تشریحی می‌تواند به‌عنوان یکی از شاخص‌های کیفیت ارزشیابی مورد استفاده قرار گیرد(۵). توزیع نرمال نشان‌دهنده پراکندگی متوازن نمرات حول میانگین است و بیان‌گر طراحی مناسب آزمون و نمره‌دهی منصفانه است. در صورت انحراف از این الگو، احتمال وجود سوگیری در تصحیح، دشواری یا سادگی بیش از حد آزمون یا عدم تفکیک مناسب بین دانشجویان وجود دارد. اگرچه نرمال بودن نمرات شرط لازم نیست، اما تحلیل آن در کنار سایر معیارهای آماری می‌تواند به بهبود اعتبار و مقایسه‌پذیری ارزشیابی‌ها کمک کند(۶). در تحلیل آماری نمرات دانشجویان، توزیع نرمال به‌عنوان الگوی کلاسیک شناخته می‌شود؛ با این حال، به‌دلیل غیرمنفی بودن نمرات و تمایل آن‌ها به تجمع در بازه‌های خاص، توزیع لگاریتمی نرمال[1] در برخی موارد نمای دقیق‌تری از پراکندگی داده‌ها ارائه می‌دهد. توزیع لگاریتمی نرمال برای داده‌های مثبت با چولگی مناسب‌تر است و می‌تواند در آزمون‌های دشوار یا با نمره‌دهی سخت‌گیرانه، ساختار واقعی‌تری را منعکس کند(۷). در این معیار ابتدا توزیع نرمال نمرات بررسی می‌شود. اگر نمرات، توزیع نرمال نداشتند، لگاریتم نمرات محاسبه شده و سپس نرمال بودن نمرات لگاریتمی بررسی می‌شود. امتیازدهی در این معیار به این صورت خواهد بود که نمرات دارای توزیع نرمال امتیاز معادل ۱، نمرات دارای توزیع لگاریتمی نرمال، امتیاز معادل

[1] Log–normal Distribution

[Downloaded from ijme.mui.ac.ir on 2026-01-30]

۱/۵ و در غیر این صورت امتیاز ۰ لحاظ خواهد شد.

- معنادار بودن اختلاف میانگین در دو گروه نمرات کم‌تر و بیش‌تر از میانگین

در حالت ایده‌آل، مطلوب است که توزیع نمرات تک‌وجهی با میانگین نمرات معقول و واریانس نسبتاً کم مشاهده شود که نشان می‌دهد اکثر کلاس به تسلط بر درس، دست یافته‌اند. اما این برداشت می‌تواند غیرواقعی باشد، زیرا کلاس به‌طور عموم به حداقل دو گروه متمایز از دانش‌آموزان با عملکرد بالا و پایین تقسیم می‌شود(۸). این معیار بررسی می‌کند که آیا تفاوت معناداری در میانگین نمرات دانشجویانی که نمرات کم‌تر و بیش‌تر از میانگین دارند وجود دارد یا خیر. اگر نمرات دانشجویان، توزیع نرمال را دنبال کند، آزمون t-test برای مقایسه میانگین دو گروه مستقل استفاده شود. اگر نمرات توزیع نرمال نداشته باشند، از آزمون غیرپارامتریک من-ویتنی برای مقایسه دو گروه مستقل استفاده شود. اگر اختلاف معنادار بین دو گروه وجود داشته باشد امتیاز این معیار ۱ و اگر وجود نداشته باشد امتیاز این معیار ۰ خواهد بود.

کلید واژه‌ها: آزمون‌های پیشرفت تحصیلی، ارزشیابی آموزشی، تحلیل آماری نمرات

منابع

1. Shakurnia A, Alijani H, Najjar S, komeili H, Elhampour H. [The Effect of Two Assessment Methods on Exam Preparation and Study Strategies: Multiple Choice and Essay Questions]. Iranian Journal of Medical Education 2013; 13 (4): 306-318.[Persian]
2. Zafaripour T, Madani SA. [Examining The Effect of Course Content Difficulty On Students' Evaluation Scores of Professors Teaching Quality]. Higher Education Letter. 2025; 18(71): 139-164. [Persian]
3. Moradi E, Didehban H. Scoring in The Essay Tests Questions: Methods, Challenges and Strategies. Nursing And Midwifery Journal 2015; 13(8): 692-698.[Persian]
4. Mohamed Ahmed IA, Moalwi AA. Correlation between Difficulty and Discrimination Indices of MCQs Type A in Formative Exam in Anatomy. IOSR Journal of Research & Method in Education. 2017; 7(5): 28–43.
5. Kulick G, Wright R. The Impact of Grading on the Curve: A Simulation Analysis. The International Journal for the Scholarship of Teaching and Learning. 2008; 2(2): 1-17.
6. Zuo J. Standard normal distribution test of college students' test score distribution. Proceedings of the 5th International Conference on Economics, Management Engineering and Education Technology (ICEMEET 2021); 2021 Jun 4-6; Toronto, Canada: 475–80.
7. Arthurs N, Stenhaug B, Karayev S, Piec C. Grades are not Normal: Improving Exam Score Models Using the Logit-Normal Distribution; 2019. [Cited 2025 Nov 5]. available from: https://stanford.edu/~cpiech/bio/papers/gradesAreNotNormal.pdf
8. Lewin DR. What Can We Learn from Exam Grade Distributions? Int J Scholarsh Teach Learn. 2021; 15(2): 1-12.

[Downloaded from ijme.mui.ac.ir on 2026-01-30]