

نقد اصول اساسی کاربرد فرمول حذف عامل شانس (با استفاده از نمره منفی) در آزمون‌ها

دکتر غلامرضا گل محمدنژاد بهرامی*

چکیده

استفاده از نمره منفی در آزمون‌های انتخاب - پاسخ یکی از مسائلی است که استفاده‌کنندگان آزمون‌ها به خصوص معلمان به طور مستمر با آن روبه‌رو هستند. فرمول حذف عامل شانس (نمره منفی) کاملاً بر اساس قانون احتمالات استوار بوده و شرایط به‌خصوصی برای استفاده از آن وجود دارد، اما در بسیاری از موارد، افراد مذکور بدون اینکه از شرایط اساسی آن آگاهی لازم را داشته باشند، آن را مورد استفاده قرار می‌دهند. زمانی که شرایط استفاده از نمره منفی محقق نشده باشد، اعمال آن جهت تعدیل نمرات، نه تنها موجب افزایش اعتبار اندازه‌گیری نخواهد شد، بلکه ممکن است نتیجه عکس داشته و آن را خدشه‌دار سازد؛ همچنین ممکن است موجب ضایع شدن حق افراد، به خصوص افراد دارای سطوح بالای توانایی شود. از طرف دیگر، این احتمال نیز وجود دارد که آزمودنی‌ها از ترس نمره منفی نتوانند آن گونه که باید و شاید توانایی واقعی خود را نشان دهند. این مقاله در نظر دارد، ضمن نقد اصول اساسی استفاده از فرمول حذف عامل شانس (نمره منفی)، استفاده‌کنندگان از آزمون‌ها را برای آگاهی از شرایط استفاده صحیح از آن یاری کند.

واژگان کلیدی: آزمون‌های انتخاب - پاسخ، آزمون‌های چند گزینه‌ای، آزمون‌های دو گزینه‌ای، عامل حدس و گمان، عامل شانس، نمره منفی.

مقدمه

در مسائل انسانی به خصوص در حوزه علوم اجتماعی، تربیتی و روان‌شناسی، «آزمون» یکی از رایج‌ترین ابزارهای اندازه‌گیری به شمار می‌آید. بیشترین استفاده از آزمون در مراکز آموزشی همانند مدارس، دانشگاه‌ها، مراکز درمانی، مشاوره و یا پژوهشی است. اهداف کاربردی آزمون‌ها گوناگون است. گاهی از آزمون‌ها جهت انتخاب و گزینش افراد، گاهی برای ارزشیابی و یا طبقه‌بندی افراد و گاهی اوقات نیز برای تشخیص بیماری‌ها و مانند آن استفاده می‌شود.

اجراء، نمره‌گذاری و تفسیر نتایج آزمون‌ها، نیازمند تخصص است. عدم تخصص لازم و استفاده نابجا از آزمون‌ها موجب زیان‌های جبران‌ناپذیری می‌شود؛ به همین دلیل، مطالعه و گذراندن دوره‌های آموزشی لازم در زمینه استفاده از آزمون‌ها ضروری به نظر می‌رسد.

آزمون‌ها دارای تقسیم‌بندی‌های مختلفی است که یکی از این تقسیم‌بندی‌ها، تفکیک آنها به آزمون‌های ساخته پاسخ^۱ و گزیده پاسخ^۲ است. آزمون‌های ساخته پاسخ، آزمون‌هایی هستند که در آنها آزمودنی‌ها خود به تولید و ساخت پاسخ سؤال‌ها مبادرت می‌ورزند؛ مانند آزمون‌های تشریحی، اما آزمون‌های گزیده پاسخ، آزمون‌هایی هستند که به آزمودنی‌ها گزینه‌هایی پیشنهاد می‌شود تا از میان آنها یکی را به عنوان درست‌ترین پاسخ انتخاب کنند؛ مانند آزمون‌های چند گزینه‌ای و یا صحیح - غلط.

تقسیم‌بندی دیگری که از آزمون‌ها وجود دارد، تقسیم آنها به آزمون‌های شناختی و غیرشناختی است. آزمون‌های شناختی شامل آزمون‌هایی است که معلومات، پیشرفت تحصیلی، استعداد، هوش و مانند آن را اندازه‌گیری می‌کنند؛ درحالی‌که آزمون‌های غیرشناختی شامل آزمون‌هایی هستند که شخصیت و یا خصوصیات روانی را می‌سنجند. در آزمون‌های شناختی گزیده پاسخ به خصوص در نوع سرعت^۳، مسئله استفاده از شانس و انتخاب تصادفی گزینه‌ها، یکی از مسائل عمده در استفاده از این نوع آزمون‌ها بوده و همواره از دغدغه‌های کاربران آنهاست؛ بدین معنی که در این نوع آزمون‌ها احتمال یافتن گزینه صحیح به صورت شانسی از طرف آزمودنی‌ها وجود دارد و ممکن

1. Constructed-response

2. Selected-response

۳. آزمون‌های سرعت، آزمون‌هایی هستند که در آن، آزمودنی‌ها باید در یک زمان محدود به تعداد زیادی از سؤال‌ها پاسخ دهند که معمولاً کسی فرصت پاسخ به همه سؤال‌ها را پیدا نکرده و تعدادی از سؤال‌ها بدون پاسخ می‌مانند. آزمون‌های سرعت در مقابل آزمون‌های قدرت طبقه‌بندی می‌شوند. در آزمون‌های قدرت تعداد سؤال محدود بوده و زمان پاسخ‌گویی نیز به حد کافی در اختیار آزمودنی‌ها قرار داده می‌شود.

است همه یا قسمتی از نمرات آنها از طریق پاسخ‌های شانسی به دست آید که این مسئله به عنوان یکی از محدودیت‌های استفاده از این آزمون‌ها مطرح است. همزمان با کاربرد و رواج تست‌های چند گزینه‌ای (گزیده - پاسخ) در دهه ۱۹۲۰، یک جریان منظمی از مطالعات پژوهشی جهت پیدا کردن راه‌حلهایی برای کاهش تأثیر عامل حدس و گمان به وجود آمد (چاپینگ، ۱۹۸۸). اهمیت عمده توجه به مسئله حدس و گمان در تست‌های چندگزینه‌ای به سه عامل جدا ارتباط دارد:

۱. عمل حدس زدن با به وجود آوردن یک خطای تصادفی باعث پایین آمدن اعتبار و روایی تست‌ها می‌شود؛
 ۲. عامل حدس و گمان باعث به دست آوردن نمرات غیرواقعی از طرف افراد و در نتیجه عدم تشخیص درست سطح پیشرفت آنها می‌شود؛
 ۳. حق کسانی که در پاسخگویی به سؤالات از حدس و گمان استفاده نمی‌کنند در مقابل کسانی که استفاده می‌کنند ضایع می‌شود (همان منبع).
- ناگفته نماند که اصطلاح «شانس» با «حدس» تفاوت اساسی دارد؛ بدین معنی که شانس پایه تصادفی داشته و کاملاً نامنظم است، اما در عین حال، از قانون احتمالات تبعیت می‌کند و شکل توزیع آن معین و مشخص است؛ درحالی که حدس و گمان همراه با آگاهی بوده و توزیع معینی ندارد. بنا به گفته «لرد» ممکن است برخی آزمودنی‌ها اطلاع کامل از جواب سؤال نداشته باشند، اما به طور کلی بعضی از آنها اطلاعات جزئی در مورد آن سؤال دارند. آنها در اکثر تست‌های چندگزینه‌ای یک یا چند گزینه را با اطمینان کمتر و یا بیشتر حذف می‌کنند و بعد یکی از گزینه‌ها را به عنوان پاسخ صحیح حدس و انتخاب می‌کنند؛ بنابراین، فرض اینکه همه آزمودنی‌ها از شانس و تصادف استفاده می‌کنند، نمی‌تواند فرض درستی باشد (لرد، ۱۹۷۵ ص ۷). با توجه به این مطالب اصطلاحی که در این بحث مد نظر بوده اصطلاح شانس و تصادف است.

اصطلاح نمره منفی که در میان عموم رایج است، همان میزان پاسخ‌های درست شانسی افراد شرکت کننده در امتحانات است که به وسیله فرمول حذف عامل شانس محاسبه و از مجموع جواب‌های درست افراد کسر می‌شود. اغلب افراد کسر این نمره را ظلمی در حق خود تلقی کرده و تصور می‌کنند که حقی از آنها ضایع شده است. به همین دلیل، دیدگاه منفی به آن داشته و با اصطلاح منفی، آن را می‌شناسند؛ در حالی که اگر اطلاع کافی از اساس آن داشته باشند ممکن

است تا این حد در مقابل آن موضع منفی اتخاذ نکنند و شاید هم به جای اصطلاح منفی، اصطلاح دیگری برای آن جایگزین کنند.

نمره منفی در اغلب امتحانات ورودی دانشگاه‌ها (کنکور) اعمال می‌شود. گاهی اوقات بعضی از استادان، دبیران و معلمان و درکل کسانی که با اجرای امتحان سر و کار دارند نیز در نمره‌گذاری اوراق آزمون‌های گزیده پاسخ (انتخاب پاسخ) نمره منفی منظور می‌کنند (مانند امتحانات پیشرفت تحصیلی در مدارس و دانشگاه‌ها) اما همین افراد و دست‌اندرکاران اغلب در اصول اساسی استفاده از نمره منفی و اینکه حتی از چه راهی نمره منفی محاسبه می‌شود و اینکه آیا برای آزمون‌هایی با گزینه‌های مختلف میزان نمره منفی متفاوت است یا نه نیز بی‌خبرند.

در این مقاله، سعی شده به نقد اصول پایه و اساسی استفاده از فرمول حذف عامل شانس و نمره منفی که اغلب افراد با فلسفه استفاده از آن و اصول و شرایط کاربرد آن آشنایی چندانی ندارند و بسیاری آن را به صورت نابجا به کار می‌برند، پرداخته شود و هر چند به اختصار، اطلاعاتی در زمینه به‌کارگیری صحیح فرمول حذف عامل شانس داده شود.

فرمول حذف عامل شانس و تصادف و اصول اساسی استفاده از آن

فرض کنید در یک آزمون ۶۰ سؤالی چهار گزینه‌ای، فردی به ۴۲ سؤال پاسخ درست و به ۱۲ سؤال پاسخ غلط داده و ۶ سؤال را بدون پاسخ گذاشته است. از آنجایی که آزمون از نوع گزیده پاسخ بوده و احتمال دارد این فرد از طریق شانس اقدام به پاسخگویی سؤال‌ها کرده باشد، باید عملکرد فرد مورد اصلاح قرار گرفته و تعداد پاسخ‌های درست شانس محاسبه و از کل پاسخ‌های درست او کسر شود. برای انجام این کار، فرمول حذف عامل شانس و تصادف که در زیر آمده است به کار می‌رود (دایموند^۱ و ایوانز^۲، ۱۹۷۳، ص ۱۸۱).

$$s = r - \frac{w}{k-1}$$

در این فرمول S برابر نمره نهایی فرد پس از کسر نمره منفی، I تعداد پاسخ‌های درست و W تعداد پاسخ‌های غلط و k تعداد گزینه‌هاست. تعداد پاسخ‌های درست نهایی فرد مذکور پس از اصلاح و کسر تعداد پاسخ‌های درست شانس (نمره منفی) برابر است با:

$$s = 42 - \frac{12}{4-1} = 38$$

1. Diamond, J.
2. Evans

برای استفاده از فرمول حذف عامل شانس دو پیش فرض اساسی وجود دارد: پیش فرض اساسی اول که در استفاده از این فرمول وجود دارد این است که فرض می‌شود کسانی که در باره پاسخ صحیح سؤال معلومات کافی ندارند آن را به صورت شانسی پاسخ می‌دهند و پاسخ‌های شانسی (تصادفی) تابع قانون احتمالات است و می‌توان از روی پاسخ‌های غلط تعداد پاسخ‌های درست شانسی را برآورد کرد؛ بدین معنی که وقتی از عامل شانس برای پاسخ به سؤال‌ها استفاده شود، تعدادی از سؤال‌ها درست و تعدادی غلط از آب در می‌آید. این استدلال بر پایه قانون احتمالات صورت می‌گیرد؛ به عنوان مثال، بر طبق قانون احتمالات اگر یک سکه سالم بی‌نهایت بار پرتاب شود نصف پرتاب‌ها به رو و نصف دیگر به پشت خواهد افتاد. شاید در ده یا بیست بار پرتاب این نتیجه حاصل نشود اما در بی‌نهایت بار پرتاب، $\frac{1}{2}$ پرتاب‌ها به رو و $\frac{1}{2}$ آنها به پشت خواهد بود. بر این اساس، اگر فردی بخواهد سؤال‌های چهار گزینه‌ای تقریباً زیادی را به طور شانسی پاسخ دهد، به احتمال فراوان $\frac{1}{4}$ پاسخ‌ها درست خواهد بود؛ برای مثال، در مورد شخصی که از ۶۰ سؤال چهارگزینه‌ای فقط به سؤال ۱۵ جواب درست داده، می‌توان استدلال کرد که وی همه سؤال‌ها را از روی شانس و تصادف جواب داده است؛ چرا که در هر سؤال چهارگزینه‌ای، شانس اینکه گزینه علامت زده شده گزینه درست باشد، فقط $\frac{1}{4}$ یا ۲۵٪ است و اگر فرد همه سؤال‌ها را تصادفی پاسخ دهد، طبق قانون احتمالات $\frac{1}{4}$ کل این پاسخ‌ها درست و $\frac{3}{4}$ دیگر غلط خواهد بود. به تبعیت از این قانون، اگر کسی به فرض ۳ جواب غلط داشته باشد می‌توان گفت احتمالاً این فرد ۴ سؤال را به طور شانسی علامت زده که پاسخ یکی از این سؤال‌ها درست و پاسخ ۳ سؤال دیگر نیز نادرست در آمده است. اگر به مثال ابتدای مطالب توجه کنید که فرد مذکور در آن از ۶۰ سؤال چهارگزینه‌ای به ۴۲ سؤال پاسخ درست داده و ۱۲ پاسخ غلط داشته، می‌توان از روی این ۱۲ پاسخ غلط استدلال کرد که وی ۱۶ سؤال را به طور تصادفی و شانسی جواب داده که $\frac{3}{4}$ یا ۷۵ درصد آنها یعنی ۱۲ پاسخ نادرست و $\frac{1}{4}$ یعنی ۴ پاسخ نیز درست از آب درآمده است و از آنجایی که باید نمرات، گویای معلومات واقعی افراد باشد، این ۴ پاسخ درست

شانسی به وسیله فرمول حذف عامل شانس از تعداد پاسخ‌های درست فرد کسر شد و در نهایت تعداد پاسخ‌های درست وی به ۳۸ رسید.

درباره این پیش‌فرض که بیان می‌کند کسانی که اطلاع کافی از پاسخ درست ندارند، اقدام به پاسخ شانس می‌کنند و باید با کاربرد نمره منفی از این کار جلوگیری کرد. این نکته به نظر می‌رسد که اولاً همه افراد شهامت یکسانی برای دادن پاسخ شانس ندارند و حتی بعضی از آنها آن را بی‌ارزش هم می‌دانند؛ ثانیاً زمانی قانون احتمالات تحقق پیدا خواهد کرد که به تعداد کافی سؤال وجود داشته باشد. اگر پاسخ به یک سؤال یک متغیر دو جمله‌ای^۱ (درست و نادرست) در نظر گرفته شود و احتمال پاسخ درست شانس با p و احتمال پاسخ نادرست شانس نیز با $q=1-p$ نشان داده شود، آن وقت به صورت نظری در یک سؤال دو گزینه‌ای $p=0/5$ و $q=1-0/5=0/5$ و در سؤال‌های چهارگزینه‌ای نیز $p=0/25$ و $q=0/75$ خواهد بود. این احتمال در بی‌نهایت بار تکرار آزمایش به دست می‌آید و برای اینکه احتمال واقعی (عملی) به احتمال نظری نزدیک شود تعداد سؤال باید به حد کافی برسد. با توجه به این مطالب، به نظر می‌رسد، استفاده از این فرمول در مورد آزمون‌هایی که تعداد سؤال‌ها کمی دارند و یا افرادی که تعداد پاسخ‌های آنها کم است، درست نخواهد بود. نکته مورد بحث دیگری که در مورد استفاده از فرمول حذف عامل شانس و نمره منفی وجود دارد، اعمال آن در مورد همه آزمودنی‌ها و یا همه نمرات است؛ به عبارت دیگر پیش فرض استفاده از نمره منفی این است که هر کسی که پاسخ اشتباهی داشته از شانس و تصادف استفاده کرده است؛ در حالی که امکان دارد همه آزمودنی‌ها از شانس و تصادف استفاده نکرده باشند؛ به عنوان مثال، کسانی که در یک آزمون نمرات بالا می‌گیرند احتمال کمی وجود دارد که از شانس استفاده کرده باشند؛ چون احتمال کسب نمرات بالا در یک آزمون به صورت شانس ضعیف است؛ به طور مثال؛ کسب نمراتی مانند ۵۷، ۵۸، ۵۹ و یا حتی ۳۸، ۳۹، ۴۰ و مانند آن به صورت شانس در یک آزمون چهار گزینه‌ای ۶۰ سؤالی خیلی ضعیف است؛ بنابراین، پیش فرض اول استفاده از فرمول حذف عامل شانس که بیان می‌کرد افراد بی‌اطلاع به صورت تصادفی و شانس پاسخ درست را پیدا می‌کنند و وجود پاسخ‌های غلط نشان‌دهنده استفاده از عامل شانس و تصادف است، در مورد افرادی که نمرات خیلی بالا گرفته‌اند صدق نمی‌کند؛ زیرا این افراد، این نمرات را شانس به دست نمی‌آورند بلکه با استفاده از دانش و معلومات خود آن را کسب می‌کنند؛ البته ممکن است این افراد هم دچار اشتباهاتی شوند و پاسخ‌های غلط نیز داشته باشند اما هر پاسخ اشتباهی دلیل استفاده از شانس و تصادف نیست؛ بنابراین، نباید نمره منفی در مورد نمرات بالا اعمال شود.

حداکثر گرفتن آن نمره به طور شانسی، ۰/۰۵ و یا کمتر است. بدین صورت در بالا و پایین نمرات دو منطقه ایجاد می‌شود که احتمال قرار گرفتن در هر یک از آنها حداکثر ۰/۰۵ است؛ برای مثال، در یک آزمون دو گزینه‌ای ۳۶ سؤالی، احتمال کسب شانسی نمرات که از فرمول بالا محاسبه شده است، به صورت زیر خواهد بود:

$$P_{(36)} = 0/0000000000145519150$$

$$P_{(35)} = 0/000000000052386894$$

$$P_{(34)} = 0/0000000009$$

$$P_{(33)} = 0/0000000103$$

$$P_{(32)} = 0/0000000857$$

$$P_{(31)} = 0/00000005485$$

$$P_{(30)} = 0/00000028344$$

$$P_{(29)} = 0/000000121474$$

$$P_{(28)} = 0/000000440345$$

$$P_{(27)} = 0/0000001369965$$

$$P_{(26)} = 0/0000003698903$$

$$P_{(25)} = 0/0000008742863$$

$$P_{(24)} = 0/00000018214261$$

$$P_{(23)} = 0/0000003362328$$

$$P_{(22)} = 0/00000055243253$$

$$P_{(21)} = 0/00000081022111$$

$$P_{(20)} = 0/00000010633036$$

$$P_{(19)} = 0/000000125103572$$

$$P_{(18)} = 0/000000132060599$$

$$P_{(17)} = 0/000000125103572$$

$$P_{(16)} = 0/00000010633036$$

$$P_{(15)} = 0/00000081022111$$

$$P_{(14)} = 0/00000055243253$$

$$P_{(13)} = 0/0000003362328$$

$$P_{(12)} = 0/00000028344$$

$$P_{(11)} = 0/000000121474$$

$$P_{(10)} = 0/000000440345$$

$$P_{(9)} = 0/0000001369965$$

$$P_{(8)} = 0/0000003698903$$

$$P_{(7)} = 0/0000008742863$$

$$P_{(6)} = 0/00000018214261$$

$$P_{(5)} = 0/0000003362328$$

$$P_{(4)} = 0/00000055243253$$

$$P_{(3)} = 0/00000081022111$$

$$P_{(2)} = 0/00000010633036$$

$$P_{(1)} = 0/000000125103572$$

$$P_{(0)} = 0/000000132060599$$

مجموع احتمال از
نمره صفر تا ۱۲ برابر
است با: ۰/۰۳۲۶۳۶۴

مجموع احتمال از
نمره ۳۶ تا ۲۴ برابر
است با: ۰/۰۳۲۶۳۶۴

در این مثال، احتمال اینکه کسی حداقل نمره ۲۴ بگیرد (یعنی ۲۴ پاسخ درست داشته باشد) ۰/۰۳۲۶۳۶۴ است که اگر نمره ۲۳ نیز به آن اضافه شود (یعنی نمره ۲۳ نیز در این محدوده قرار گیرد) احتمال گرفتن حداقل ۲۳ بیش از ۰/۰۵ خواهد بود که در

جامعه علمی قبول ریسک بیش از ۰/۰۵ جایز نیست و همین طور در منطقه پایین احتمال گرفتن حداکثر نمره ۱۲ (یعنی داشتن حداکثر ۱۲ پاسخ درست) باز مساوی ۰/۰۳۲۶۳۶۴ است؛ بنابراین، در اینجا دو محدوده بالا و پایین ایجاد شده است که احتمال قرار گرفتن در هر یک از آنها به تنهایی کمتر از ۰/۰۵ است و چون افراد در این محدوده‌ها قرار بگیرند با ریسک کمتر از ۰/۰۵ می‌توان گفت که آنها به صورت تصادفی و شانسی در این محدوده‌ها قرار نگرفته‌اند و دلیلی که باعث شده این افراد در این محدوده‌ها باشند، دانش و معلومات آنان بوده است.

برخی آماردانان از جمله «توایت^۱ و مونرو^۲» (به نقل از هومن، ۱۳۷۰) پیشنهاد کرده‌اند در مسائل عملی، هنگامی که مقدار $n \cdot pq$ حداقل برابر با ۹ باشد ($n \cdot pq \geq 9$) توزیع متغیر دو جمله‌ای به توزیع نرمال^۳ نزدیک می‌شود؛ بنابراین، می‌توان از مقادیر توزیع نرمال برای آزمون فرضیه‌های آماری و یا برآوردهای مختلف استفاده به عمل آورد.

با توجه به مطلب بالا، در صورتی که آزمون‌های دو گزینه‌ای یا صحیح- غلط دارای حداقل ۳۶ سؤال

$$\{(n \times 0.5 \times 0.5 \geq 9) \Rightarrow (n \times 0.25 \geq 9) \Rightarrow (n \geq \frac{9}{0.25}) \Rightarrow (n \geq 36)\}$$

و آزمون‌های چهارگزینه‌ای دارای حداقل ۴۸ سؤال

$$\{(n \times 0.25 \times 0.75 \geq 9) \Rightarrow (n \times 0.1875 \geq 9) \Rightarrow (n \geq \frac{9}{0.1875}) \Rightarrow (n \geq 48)\}$$

باشند، می‌توان توزیع مربوط به احتمال نمرات تصادفی را تقریباً نرمال فرض کرد؛ در صورتی که توزیع دو جمله‌ای به توزیع نرمال نزدیک شود با استفاده از فرمول زیر می‌توان محدوده ۵ درصدی (۰/۰۵) در دو انتهای توزیع و یا محدوده ۹۰ درصدی (۰/۹) در وسط توزیع را مشخص کرد که در مقایسه با روش قبلی که در آن احتمال هر کدام از نمرات جداگانه و در دو قسمت بالا و پایین توزیع محاسبه و سپس جمع می‌شدند، روش آسانی به حساب می‌آید.

$$\pm (n \cdot p) \pm (\sqrt{n \cdot pq} \times 1.65)$$

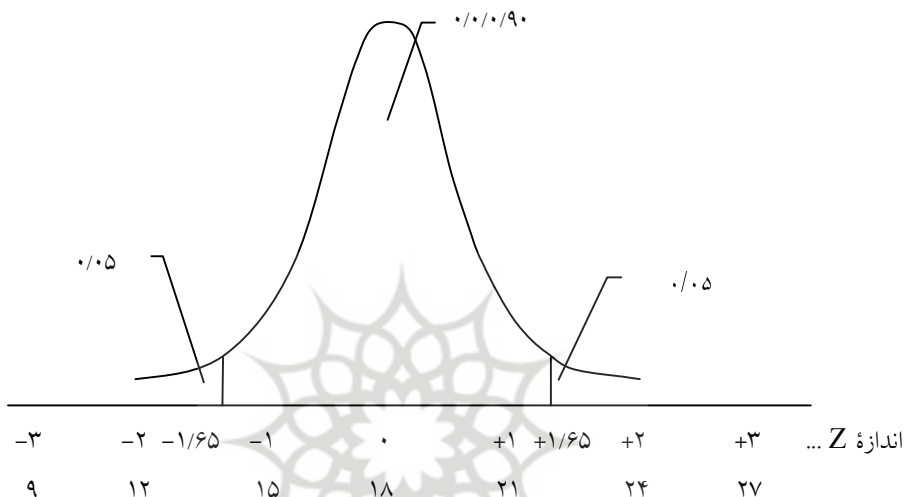
در این فرمول، n تعداد سؤال‌ها آزمون، p و q هم به ترتیب احتمال موفقیت و عدم موفقیت به صورت تصادفی در هر سؤال است.

1. Twaite
2. Monroe
3. Normal distribution

۴. این فرمول برای تعیین محدوده ۹۰ درصدی بر اساس تقریب نرمال دو جمله‌ای از سوی نویسنده مقاله از منابع آمار استنباطی استخراج شده است و در منبع خاصی عیناً نیامده است.

اگر توزیع احتمال نمرات تصادفی یک آزمون ۳۶ سؤالی دو گزینه‌ای نرمال فرض شود، محدوده ۹۰ درصدی میانی توزیع و یا ۵ درصدی در دو سوی انتهایی توزیع با استفاده از فرمول مذکور چنین خواهد بود:

$$(36 \times 0.5) \pm \left[\sqrt{(36 \times 0.25) \times 1/65} \right] < \begin{matrix} 22/95 \approx 23 \\ 13/05 \approx 13 \end{matrix}$$



با در نظر گرفتن حدود واقعی^۱ نمرات، احتمال به دست آوردن تصادفی نمره‌ای بین ۱۲/۵ و ۲۳/۵، ۰/۹ و احتمال کسب نمراتی بالاتر از ۲۳/۵ و یا کمتر از ۱۲/۵ نیز هر کدام جداگانه ۰/۰۵ خواهد بود. علوم انسانی و مطالعات فرهنگی

اینک در این مثال، چون ما نمی‌توانیم گرفتن نمرات بالاتر از ۲۳ و پایین‌تر از ۱۳ را به شانس و تصادف نسبت دهیم (با ریسک ۰/۰۵)، جایز نیست قانون نمره منفی را در باره آنها اعمال کنیم؛ زیرا نمره منفی مخصوص کسانی است که به صورت شانسی و تصادفی نمره گرفته‌اند؛ البته استفاده از فرمول مذکور برای تعیین محدوده ۹۰ درصدی توزیع نمرات تصادفی، محدود به سؤال‌های دو گزینه‌ای و یا صحیح - غلط نبوده و در مورد هر سؤالی با هر تعداد گزینه از جمله سؤال‌های چهارگزینه‌ای می‌توان استفاده کرد، اما پیش فرض دومی که استفاده از فرمول حذف عامل شانس بر

۱. حدود واقعی هر نمره، نیم واحد اندازه‌گیری بالاتر (حد بالا) و نیم واحد اندازه‌گیری پایین‌تر از نمره (حد پایین) را شامل می‌شود؛ به عنوان مثال، اگر واحد اندازه‌گیری ۱ باشد، حدود واقعی نمره ۱۳ بین نمره ۱۲/۵ و ۱۳/۵ خواهد بود.

آن استوار است، عبارت است از اینکه گزینه‌های یک سؤال همگی چه درست و چه نادرست دارای جاذبه یکسان برای افراد بی اطلاع هستند؛ یعنی احتمال انتخاب شدن هر یک از گزینه‌ها نزد افراد کم اطلاع با همدیگر مساوی است. در مورد پیش فرض دوم نیز این ایراد وجود دارد که اگرچه ساختن سؤالی که همه گزینه‌های آن جاذبه یکسانی داشته باشند، محال نیست اما به نظر می‌رسد، در عمل بیش از حد دشوار و غیر عملی باشد که حتی از عهده متخصصان سنجش و اندازه‌گیری هم خارج است؛ در صورتی که گزینه‌های سؤال برای افراد بی اطلاع جاذبه یکسانی نداشته باشد، پاسخ‌های شانس افراد نیز از قاعده و اصول احتمالات پیروی نخواهد کرد.

نتیجه‌گیری

در مورد استفاده از فرمول حذف عامل شانس در آزمون‌های چند گزینه‌ای و صحیح - غلط علاوه بر بحث مربوط به تحقق فرضیه‌های زیربنایی فرمول مذکور، بحث دیگری نیز به تأثیر آن در اعتبار و روایی آزمون و مفید بودن و یا نبودن استفاده از آن مربوط است.

«یولا بربی»^۱ طی مقاله‌ای تحقیقی عامل حدس و گمان را باعث افزایش واریانس خطا در نمرات و در نتیجه پایین آمدن اعتبار آزمون‌ها قلمداد کرده و در تحقیقاتی که «بارتون»^۲ (۲۰۰۱) و «پریهدا»^۳ و همکارانش (۲۰۰۶) در مورد تأثیر استفاده از فرمول حذف عامل شانس بر اعتبار و روایی آزمون‌های چند گزینه‌ای انجام داده‌اند، نتایج به دست آمده حاکی از افزایش اعتبار و روایی این آزمون‌ها به هنگام استفاده از فرمول مذکور بوده است.

مطلب مهمی که در مورد تحقیقات بالا وجود دارد، این است که هر چند این تحقیقات تأثیر استفاده از فرمول حذف عامل حدس و گمان را در افزایش اعتبار و روایی آزمون‌ها نشان می‌دهند اما این تأثیر چندان هم که باید چشمگیر نیست؛ همچنین اغلب این تحقیقات در مورد آزمون‌های دو گزینه‌ای انجام شده که شانس افراد برای یافتن پاسخ صحیح در آنها بیشتر است.

به رغم تحقیقات مذکور، بررسی‌هایی نیز نشان داده که وقتی تعداد گزینه‌ها افزایش می‌یابد (۴ و یا بیشتر) و در مورد آزمون‌های قدرت که در آن فرصت کافی در اختیار آزمودنی‌ها قرار داده می‌شود، اعمال فرمول حذف عامل حدس و گمان چندان تأثیری در نمرات افراد ندارد و در واقع بین نمرات به دست آمده از فرمول تصحیح و

1. Bereby Yoella
2. Burton Richard F.
3. Prihoda Thomas J.

نمرات جمع ساده پاسخ‌های درست، یک رابطه خطی و مستقیم وجود دارد (هومن، ۱۳۷۲).

به طور کلی عامل شانس و تصادف موقعی تأثیر جدی در امر نمره‌گذاری و در نهایت سنجش و اندازه‌گیری دارد که آزمون دارای سؤال‌های دو گزینه‌ای بوده (که در آن احتمال پاسخ درست شانسی ۵۰ درصد است) و یا از نوع سرعت باشد. در آزمون‌های سرعت، ممکن است بخشی از آزمودنی‌ها به دلیل محدود بودن زمان، سؤال‌ها را با سرعت و تصادفی پاسخ دهند، و عده‌ای دیگر نیز در حالی که تعدادی سؤال بی‌پاسخ دارند استفاده از شانس را بی‌ارزش تلقی کنند و نخواهند سؤال‌های بی‌پاسخ خود را به صورت تصادفی جواب دهند.

در آزمون‌های چهارگزینه‌ای ورودی دانشگاه‌ها، به دلیل محدود بودن زمان اجرای آزمون، احتمال استفاده افراد از عامل شانس برای پاسخ دادن به سؤال‌ها بیشتر بوده و از آنجایی که هدف این آزمون‌ها گزینش افراد مستعد است، استفاده از نمره منفی در این آزمون‌ها موجه جلوه می‌کند، اما همان گونه که قبلاً نیز گفته شد، احتمال کمتری وجود دارد، افرادی که نمرات خیلی بالا و یا خیلی پایین کسب کرده‌اند از شانس و تصادف استفاده کرده باشند؛ به خصوص اینکه آزمون چهارگزینه‌ای بوده و احتمال موفقیت شانسی در هر سؤال فقط ۰/۲۵ است که در این شرایط احتمال کسب نمرات شانسی بسیار بالا و یا بسیار پایین ناچیز خواهد بود؛ بنابراین، در این آزمون‌ها نمره منفی تنها در مورد افرادی باید به کار گرفته شود که احتمال کسب نمره شانسی از طرف آنها بیش از ۰/۰۵ باشد؛ به عبارت دیگر، نمره منفی در مورد حدود ۹۰ درصد میانی توزیع نمرات (از طریق فرمولی که در بالا ارائه شد قابل محاسبه است) اعمال شود.

به طور خلاصه می‌توان گفت، هنگامی که تعداد سؤال زیاد بوده و زمان پاسخگویی نیز محدود باشد (آزمون‌های سرعت) و یا تعداد گزینه‌های آزمون کم باشد (همانند آزمون‌های دو گزینه‌ای) و یا هدف از به‌کاربردن آزمون گزینش افراد باشد، می‌توان از فرمول حذف عامل شانس جهت تعدیل و اصلاح عملکرد افراد استفاده کرد؛ البته در سنجش ملاک مرجع می‌توان با بالا بردن سطح ملاک چیرگی تا حدودی تأثیر استفاده از شانس را کاهش داد (مثلاً نمره قبولی افراد از ۱۰ به ۱۲ و یا بالاتر تغییر داده شود).

در خاتمه باید افزود، چنان‌که «پلوم لی»^۱ (به نقل از هومن، ۱۳۷۲) نیز اشاره می‌کند، اگر آزمونی که به‌کار برده می‌شود، دارای گزینه‌های متعدد مثلاً ۴، ۵ و یا ۶ گزینه‌ای و یا از نوع آزمون‌های قدرت محسوب شود، اقدام در مورد اعمال نمره منفی

و اصلاح نمرات فعالیت بی‌هوده است. پس مشاهده می‌شود که استفاده از فرمول تصحیح (حذف عامل شانس) به شرایط خاص که حاکم بر آزمون و چگونگی اجرای آن است، بستگی دارد؛ در غیر این صورت، استفاده از آن، در غیر از شرایط گفته شده نه تنها باعث اتلاف وقت و انرژی می‌شود (در حالی که مطلوب است به افراد اجازه حدس و گمان داده شود) بلکه ممکن است استفاده از نمره منفی موجب شود آزمودنی‌ها به خاطر ترس از آن نتوانند توانایی‌هایشان را نشان دهند و در مواقعی نیز امکان دارد کار اندازه‌گیری را مشکل و آن را بی‌اعتبار سازد؛ بنابراین، فرمول حذف عامل شانس، علاوه بر اینکه بهتر است در مورد آزمون‌های سرعت و یا دو گزینه‌ای استفاده شود، لازم است گزینه‌های سؤال‌ها درست طراحی شود و جاذبه یکسانی داشته باشند. همین‌طور باید سؤال‌ها به تعداد کافی زیاد بوده و فرمول مذکور نیز باید در مورد نمراتی به کار برده شود که احتمال کسب تصادفی آن نمرات بالا باشد.



منابع

- هومن، حیدر علی (۱۳۷۰). استنباط آماری در پژوهش رفتاری. نشر مؤلف.
- هومن، حیدر علی (۱۳۷۲). اندازه‌گیری‌های روانی و تربیتی و فن تهیه تست. نشر پارسا.
- هومن، حیدر علی (۱۳۷۰). ریاضیات پایه آمار و احتمال. نشر مؤلف.

- Bereby, Y. & Joachim, M. & Flascher, O. (2002). Prospect theory analysis of guessing in choice tests. *Journal of Behavioral Decision Making*. Vol. 15, No. 4, pp. 313-327.
- Burton, R. F. (2001). Misinformation, partial knowledge and guessing in true/false tests. *Journal of Medical Education*. Vol. 36, No.9, P. 805-811.
- Chopping, B. H. (1988). *Correction for guessing*. Educational Research, Methodology and Measurement, Pergamon Press, Plc, 384-386.
- Diamond, J. & Evans, W. (1973). The Correction for guessing. *Review of Educational Research*, Vol. 43, No.2, pp. 181-191.
- Lord, F. M. (1975). Formula scoring and number-right scoring. *Journal of Educational Measurement*, Vol. 12, No. 1.
- Prihoda, T. J. & Neal Pinckard, R. & Alex McMahan, C. & Anne Cale Jones. (2006). Correcting for guessing increases validity in multiple-choice examinations in an Oral and maxillofacial pathology course. *The Journal of Dent Educ.* 70(4), pp. 378-386.
- Plumlee, L. B. (1952). The effect of difficulty and chance success on item-test correlation and on test reliability. *Psychometrika*, 17, 69-86.
- Plumlee, L. B. (1952). The effect of difficulty and chance success on item-test correlation and on test reliability. *Psychometrika*, 19, 65-70.
- Twaite, J. A. & Monroe, J. A. (1979). *Introductory statistics*. Glenview, Illinois: Scott, Foresman and Company.