



دانشگاه علوم پزشکی بقیة الله

تجزیه و تحلیل سوال های چند گزینه ای



دانشگاه علوم پزشکی بقیة الله

MCQ Analysis

کاربرد تجزیه و تحلیل سؤال ها

- انتخاب سؤال های بهتر و مناسب تر
- درك نقاط قوت و ضعف دانشجویان
- درك نقاط قوت و ضعف تدریس و بهبود روش تدریس
- بررسی گزینه های نادرست و نامتناسب
- ترتیب سؤال ها بر اساس درجه دشواری آنها
- امکان کوتاه کردن طول تست همزمان با افزایش ضرایب اعتبار و روایی تست

تعريف

درک چگونگی تفسیر نمرات دانشجویان به
منظور کمک در جهت طراحی مطلوب سؤال های
آزمون و بهبود کیفیت آموزش.

نمره گذاری

- تصحیح دستی (سوراخ کردن جواب ها در یک پاسخ نامه)

- تصحیح ماشینی (دستگاه کد خوان)

تعیین عامل حدس وگمان (نمره منفی) :

$$R - \frac{W}{N - 1}$$

نمره اصلاح شده

R = تعداد جواب درست

W = تعداد پاسخهای غلط

N = تعداد گزینه ها

هدف از تجزیه و تحلیل نتایج امتحان

۱- هر یک از یادگیرندگان به کدامیک از اهداف آموزشی رسیده و به کدام هدف ها نرسیده است ؟

از روی پاسخ عملکرد هر فراگیر به هر یک از سوالات امتحان

۲- به طور کلی هر یک از یادگیرندگان تا چه اندازه ای به هدف های آموزشی رسیده است ؟

از روی درصد عملکرد هر فراگیر در کل آزمون

..... هدف از تجزیه و تحلیل نتایج امتحان

۳ - به طور متوسط یادگیرندگان کلاس تا چه اندازه ای به هدف های آموزشی رسیده است ؟

از روی درصد عملکرد کلاس در هر سوال

۴ - به طور کلی یادگیرندگان کلاس تا چه اندازه ای به هدف های آموزشی رسیده اند ؟

از روی درصد عملکرد کلاس در کل آزمون

- پرسش های شماره ۱ و ۲ نقاط ضعف هر فراگیر را نشان می دهد .
مدرس با آگاهی از آنها می تواند با آموزش مجدد ، آموزش جبرانی ، تمرین های لازم و ... آنها را بهبود ببخشد .
- پرسش شماره ۳ نشان دهنده نقاط قوت و ضعف تعدادی از فراگیران در هر یک از هدف های آموزشی است
که می تواند نشانگر نارسایی در مواد و وسایل آموزشی، روش تدریس مدرس و یا کیفیت سوالات امتحانی باشد .
- پرسش شماره ۴ سطح کلی موفقیت کلاس در تحقق اهداف درس را نشان می دهد در عین حال مبین سطح کارایی مدرس در امر تدریس است .

هدف از تحلیل سوال های آزمون

1. تعیین میزان دقت آزمون
2. نارسایی های آزمون
3. نقاط قوت و کیفیت سوالات بررسی می شود



تجدید نظر در آزمون و بهبود کیفیت سوالات

جدول کار فرضی (آزمون ۸ سوالی در مورد ۱۶ فراگیر)

نام فرا گیران	شماره سئوالات امتحان									جمع ردیفها	درصد عملکرد هر فراگیر در کل آزمون
	۱	۲	۳	۴	۵	۶	۷	۸			
۱-الف ۲-ب	۱ ۱	۰ ۱	۱ ۱	۱ ۱	۱ ۰	۱ ۱	۱ ۰	۰ ۰	۶ ۵	۷۵	
۱۶-ش	۰	۰	۱	۰	۰	۰	۰	۰	۱	۱۲/۵	
جمع ستونها	۱۱	۱۰	۱۶	۱۵	۷	۸	۴	۲	۷۳	۹۱۲	
درصد عملکرد کلاس در هر پرسش	۶۸/۷۵	۶۲/۵	۱۰۰	۹۳/۷۵	۴۳/۷۵	۵۰	۲۵	۱۲/۵		۵۷	درصد عملکرد کلاس در کل آزمون

پس از تهیه جدول با مراجعه به ورقه هر فراگیر در مقابل نام او در ردیف مربوط به شماره هر سوال برحسب اینکه پاسخ درست یا غلط داده باشد عدد ۱ یا صفر می‌نویسیم. پس از کامل شدن در جدول جمع ستونها و ردیفها، حاصل جمع آنها نوشته می‌شود. سپس در آخرین ستون این جدول، درصد عملکرد هر فراگیر در کل آزمون نوشته می‌شود از طریق

$$\frac{\text{تعداد پاسخهای درست هر فراگیر}}{\text{تعداد سوالهای امتحان}} \times 100$$

و در آخرین ردیف جدول از پائین درصد عملکرد کلاس در هر پرسش نوشته می‌شود از طریق

$$\frac{\text{تعداد کسانی که به هر پرسش پاسخ درست داده‌اند}}{\text{تعداد فراگیران کلاس}} \times 100$$

و در آخرین خانه ستون آخر، درصد عملکرد کلاس در کل آزمون نوشته می‌شود از طریق

$$\frac{\text{مجموع اعداد مندرج در آخرین ستون جدول}}{\text{تعداد آنها}} \times 100$$

مشخصه های تجزیه و تحلیل سؤال

1. دشواری سؤال (P)

2. قدرت تشخیص سؤال (D)

3. پایایی آزمون Reliability

4. ارزشیابی گزینه های سؤال

شیوه تجزیه و تحلیل سوال های آزمون

- امتحان را برگزار نمایید.
- اوراق دانشجویان را تصحیح و نمره گذاری نمایید.
- اوراق صحیح شده را به ترتیب نمره کسب شده از بالا به پایین مرتب کنید.
- گروه قوی و ضعیف را جدا کنید (27% از بالا و 27% از پایین).
- اطلاعات بدست آمده را در جدول تجزیه و تحلیل سوال ها وارد کنید و شاخص های مربوط به هر سوال را محاسبه کنید.

مراحل تحلیل سوالات

• تهیه کارت تحلیل سوال

1. در هر سوال چند نفر گزینه های درست را انتخاب کرده اند؟
2. هر یک از گزینه های انحرافی چند نفر را به خود جلب کرده است؟
3. چند نفر به سوال پاسخ نداده اند؟

تجزیه و تحلیل گزینه های سوالهای چند گزینه ای

- گزینه های انحرافی باید دارای 2 ویژگی باشند:
 - 1- توجه افراد بی اطلاع از موضوع را به اندازه گزینه درست به خود جلب کنند.
 - 2- برای افرادی که توانایی و مهارت کافی دارند، گول زننده نباشند.
- مرور همه گزینه های انحرافی و بررسی عملکرد آن ها می تواند منجر به اصلاح، جایگزینی یا حذف گزینه های انحرافی ضعیف شود.
- استفاده از فراوانی درصد انتخاب گزینه های انحرافی نشان دهنده نسبت دانشجویانی است که گزینه های انحرافی را انتخاب کرده اند.
- حذف یا جایگزینی گزینه های انحرافی

جدول پراکندگی پاسخ ها به سوالات آزمون

بدون جواب	د	ج	ب	الف	شماره سوال
					1
					2
					3
					4
					

در نگاه کلی به جدول پراکندگی

1. حتی المقدور توزیع گزینه ها به گونه ای باشد که بدون پاسخ نداشته باشیم یا نادر باشد.
2. اگر گزینه ای بدون پاسخ باشد به معنی این است که سوال مشکل جدی دارد.
3. اگر یکی از گزینه های نادرست بیشتر از گزینه درست باشد، باید به این موضوع فکر کنیم،

که ممکن است پاسخ غلط، درست باشد

جدول تحلیل یا توزیع پاسخ ها

نوع عکس العمل به سوال شماره	جواب درست	جواب نادرست	بدون جواب
۱	۴۵ نفر	۱۵ نفر	۵ نفر
۲	۵۲ نفر	۱۰ نفر	۳ نفر
۳	۳۸ نفر	۲۷ نفر	-
۴	۶۰ نفر	۴ نفر	۱ نفر
.....			

در نگاه کلی (بدون تحلیل) باید پاسخ های صحیح از نادرست بیشتر باشد . اگر نبود یعنی باید در سوال بازنگری کنیم

مراحل تحلیل سوالات

• تعیین نحوه پاسخ دهی آنها به گزینه های مختلف هر سوال

1. تعیین افراد گروه قوی که هر یک از گزینه های سوال را انتخاب کرده اند یا آن را بی جواب گذاشته اند.

2. تعیین افراد گروه ضعیف که هر یک از گزینه های سوال را انتخاب کرده اند یا آن را بی جواب گذاشته اند.

ضریب دشواری سوال یا ضریب آسانی یا سهولت

Difficulty Index or Facility Index

- در صد کل آزمودنی هایی که به یک سوال جواب مثبت می دهند.
-

که با حرف P نشان داده می شود .

- هر اندازه ضریب دشواری یک سوال بزرگتر باشد (یعنی به ۱۰۰ یا ۱) آن سوال آسانتر است و هرچه به سمت صفر نزدیک شود دشوارتر است .

- ضریب دشواری مطلوب بین $\frac{3}{10}$ تا $\frac{7}{10}$ است . عدد منفی بیانگر دشواری انحرافی است و باید حذف شود .

نحوه محاسبه ضریب دشواری سوال یا ضریب آسانی یا سهولت

۱- ضریب دشواری خام (ساده ترین روش)

$$P_i = r_i / n$$

تعداد افراد با پاسخ صحیح تقسیم بر کل افراد
این روش دقت زیادی ندارد ولی سهولت اندازه گیری بالایی دارد .

انتخاب درست گروه بالا + انتخاب درست گروه پایین

۲-

تعداد افراد گروه بالا + تعداد افراد گروه پایین

$$P_i =$$

.... نحوه محاسبه ضریب دشواری سوال

۳- شاخص دشواری اصلاح شده :

- در این فرمول واکنش افراد به سوالات صحیح و غلط محاسبه شده و هم تعداد گزینه هادر نظر گرفته می شود .

$$P_{ci} = \frac{R_i + W_i/K - 1}{R_i + W_i}$$

R_i = تعداد جواب های صحیح

W_i = تعداد جواب های نادرست

K = تعداد گزینه

.... نحوه محاسبه ضریب دشواری سوال

۴ - شاخص دشواری هورست :

- فراوانترین واکنش غلط را هم اضافه می کند .
- این شاخص از دقت بالاتری برخوردار است .

R_i = پاسخ های صحیح به سوال

D_i = فراوانترین پاسخ غلط

W_i = مجموع پاسخ های غلط هر سوال

$$\text{Phi} = \frac{R_i - D_i}{R_i + W_i}$$

تفسیر ضریب دشواری

- در آزمون های هنجاری، پراکندگی نمرات یا واریانس های بزرگتر نشانه مطلوبیت آزمون است .

$$P(1-P) = \text{واریانس سوال}$$

$$\text{حد مطلوب سطح دشواری هر سوال} = \frac{\text{احتمال پاسخ تصادفی} + 1}{2}$$

حد مطلوب سطح دشواری برای سوالات ص و غ ، چهار گزینه ای و کوتاه پاسخ را حساب کنید

محاسبه ضریب تمیز سوال Discrimination Index یا (ضریب تشخیص)

- تعریف: قدرت سوال را در تمایزگذاری یا تشخیص بین گروه قوی و ضعیف آزمون شوندگان مشخص می کند. که با حرف d نشان داده می شود.

$$d = \frac{\text{انتخاب درست گروه پایین} - \text{انتخاب درست گروه بالا}}{\text{تعداد افراد یک گروه (بالا یا پایین)}}$$

تفسیر ضریب تمیز (ضریب تشخیص)

1. مقدار آن بین ۰ تا ۱ متغیر است .
2. هر قدر ضریب تمیز به ۱ نزدیکتر باشد ؛ قوه تمیز آن بزرگتر است .
3. اگر ضریب تمیز منفی شد ، یعنی در این سوال گروه قوی بدتر از گروه ضعیف عمل کرده است و احتمالا این سوال اشکال فنی دارد .
4. ضریب تمیز سوال نوعی ضریب توافق (ضریب همبستگی) سوال با کل آزمون است .

رابطه بین ضریب دشواری و ضریب تمیز

- سوال های خوب در یک آزمون سوالاتی است که دارای ضریب دشواری متوسط (0/5) و ضریب تمیز بالایی (+1) است .

تجدید نظر و اصلاح آزمون

پس از تعیین ضریب های دشواری و تمیز تمامی سوالات و بررسی گزینه های هر سوال ، باید بر اساس نتایج به اصلاح یا حذف

آنها اقدام کرد .سوالاتی که ضریب دشواری بالا یا بسیار پایین و

ضریب تمیز آنها خیلی کم است باید مورد تجدید نظر قرار گیرند . همچنین گزینه های انحرافی معیوب نیز باید اصلاح یا عوض شوند .

مثال

- 60 دانشجو در آزمون شرکت داشتند. 27 درصد گروه بالا و پائینی هر کدام حدود 16 نفر می شود.

- از 16 نفر گروه بالا (قوی) 5 نفر به سوال یک پاسخ صحیح دادند و از 16 نفر گروه پایین (ضعیف) فقط یک نفر به همان سوال پاسخ صحیح داده است. بنابراین ضرایب دشواری و تمیز به شرح زیر است

$$p = \frac{5+1}{32} = .19$$

$$D = \frac{5-1}{16} = .25$$

سطح دشواري سؤال

- دشواري سؤال (P) "درصد دانشجوياني که به سؤال پاسخ درست دادند".
- دامنه دشواري بين صفر تا 1 (0% تا 100%) است.
- ضريب دشواري یک (1) يا نزديک به آن، نشان دهنده آساني سؤال است.
- ضريب دشواري بالاي 0/7 نشان دهنده سادگي بيش از حد سؤال است که نبايستي در آزمون هاي بعدي مورد استفاده قرار گيرد.
- اگر همه دانشجويان به سؤال ها پاسخ درست داده باشند محتواي سؤال مشتمل بر مفهومي است که احتمالاً براي آزمون ارزشي ندارد.
- ضريب دشواري زير 0/3 نشان دهنده سختي بسيار زياد سؤال است.
- اگر تقريباً همه دانشجويان به سؤال پاسخ غلط داده باشند، يا سؤال غلط است يا دانشجو نتوانسته مفهوم را درک کند.
- راه حل : بازنگري در کلمات گيج کننده،
- خارج کردن سؤال از آزمون هاي بعدي
- تدريس مجدد هدف آموزشي

قدرت تشخیص (تمیز) سؤال (D)

- دامنه آن از 1- تا 1+
- مقادیر بالا، نشان دهنده قدرت تشخیص بالایی سوال است.
- قدرت تشخیص بالا، نشان دهنده آن است که دانشجویانی که نمرات پایینی گرفته اند به آن سؤال پاسخ غلط داده اند.
- سؤال با قدرت تمیز زیر صفر یا نزدیک به صفر بیانگر:
 - عملکرد بهتر دانشجویان ضعیف
 - عملکرد ضعیف دانشجویان قوی.
- این سؤال دارای ابهام، گنگی و نکات گیج کننده برای دانشجویانی قوی است.

Reliability اعتبار آزمون

- اعتبار آزمون بیانگر میزان ثبات و پایایی آزمون در رابطه با نمرات آزمون است.
- دامنه آن از صفر تا یک است.
- مقادیر بالا نشان دهنده:
 - اعتبار زیاد نمرات آزمون.
 - همبستگی بالایی سؤال‌های خوب با سایر سؤال‌ها.
 - همبستگی بالا یعنی مجموعه سؤال‌ها، مفاهیم مشابه را مورد سنجش قرار می‌دهد.

RELIABILITY

TEST-RETEST

(COEFFICIENT OF STABILITY)

PARALLEL FORM

(COEFFICIENT OF EQUIVALLENCE)

INTERNAL CONSISTENCY

روشهاي بهبود اعتبار آزمون

- افزايش تعداد سؤال ها
- استفاده از سؤال هايي با قدرت تشخيص بالا

Reliability مقدار

0/9 به بالا

0/8 – 0/9

0/7 – 0/8

0/6 – 0/7

0/5 – 0/6

زیر 0/5

تفسیر

اعتبار عالی (آزمون های استاندارد)

بسیار خوب برای آزمون های پیشرفت تحصیلی

مناسب و خوب برای آزمون های پیشرفت تحصیلی
(تعدادی از سؤال ها نیاز به بهبود دارند)

تقریباً پایین
(باید بعضی سؤال ها حذف شوند)

پیشنهاد بازنگری کلی آزمون
(احتمالاً سؤال های آزمون کم بوده است)

اعتبار آزمون زیر سؤال است
(این آزمون نباید مورد استفاده قرار گیرد و نیاز به بازنگری کلی دارد)



تو خستود بایستی و مار سنگار

خایا چنان کنی سر انجام کار