

نحوه تجزیه و تحلیل نتایج آزمون

گردآورنده: ابوالقاسم رافع

دفترارزشیابی مهارت

آذرماه ۱۳۸۸

نحوه تجزیه و تحلیل نتایج آزمون

وای بر آن که نداند. و وای بر آن که بداند و عمل نکند. حضرت محمد صل الله و علیه و آله و سلم.
مردم دشمن چیزی هستند که نمی دانند. امام علی علیه السلام
کسی که محل ورود را نشناسد از یافتن محل خروج درمانده گردد. امام جواد علیه السلام
کسی که بندگی خدا کند خداوند همه چیز را بنده او کند. امام حسن مجتبی علیه السلام
افسوس کوتاهی در کارهای گذشته را با تلاش در آینده جبران کنید. امام هادی علیه السلام

مقدمه:

یک بخش از تجزیه و تحلیل نتایج آزمون این است که در پایان هر آزمون شاخص های مربوط به درجه دشواری، ضریب تمیز هر سوال، اعتبار درونی سوالات، اعتبار کل آزمون و روایی، محاسبه می شود و با نتایج تجزیه و تحلیل سوالات از قبیل درجه دشواری نظری، روایی ظاهری مقایسه می گردد تا پس از اجرای هر آزمون نسبت به اصلاح شاخص های ثبت شده در بانک سوال و حذف سوالات غیر قابل قبول و جایگزینی سوالات بهتر در بانک اقدام و تذکرات لازم به طراحان سوال ارایه شود تا افزایش کیفیت طراحی سوال موجب سنجش دقیق تر مهارت در کشور شود و هر قدر اندازه گیری دقیق تر باشد، تصمیم هایی که در مورد آزمون شوندگان و بهبود آموزش گرفته می شود به واقعیت نزدیکتر خواهد بود.

با استمرار این کار و مقایسه نتایج حاصل از تجزیه و تحلیل سوالات طرح شده در آزمونهای بعدی، میزان تاثیر آن مشخص خواهد شد.

بخش دیگر از تجزیه و تحلیل نتایج آزمون، بررسی تعداد قبول شدگان در آزمون های کتبی و عملی به تفکیک حرفه - شهرستان و جنس است. از این اقدام امکان مقایسه عملکرد آزمون شوندگان در سطح شهرستان، استان و کشور فراهم می شود. رسم نمودار و تهیه جداول آماری با استفاده از نرم افزار تصحیح پاسخنامه ها ویا excel امکان پذیر می باشد.

ابزار گردآوری داده ها و تجزیه و تحلیل نتایج آزمون:

- تکمیل چک لیست مربوط به روایی ظاهری، روایی محتوایی و درجه دشواری نظری سوالات.
- محاسبه شاخصها (مانند: درجه دشواری سوالات، ضریب تمیز سوالات، روایی و اعتبار آزمون) و مقایسه آن با نتایج شاخصهای نظری.
- تعیین و تحلیل گزینه های انحرافی.
- استفاده از ضریب همبستگی دورشته ای نقطه ای، آزمون های آماری t و آزمون های مشابه برای تجزیه و تحلیل داده ها.

هدفهای رفتاری:

- انتظار می رود هر فراگیر در پایان آموزش، مطالعه جزوه و انجام تمرین قادر باشد:
- مفاهیم ارزشیابی، اندازه گیری و آزمون را از یکدیگر بازشناسد.
- انواع ارزشیابی و ویژگی های آزمون را توضیح دهد.
- سؤالات و گزینه های مطلوب و نامطلوب را تحلیل کند.
- محاسبات مربوط به هر یک از اصطلاحات زیر را به کمک نرم افزار انجام دهد:

اعتباروروایی آزمون، درجه دشواری، ضریب تمیز هر سؤال ، میزان کارآیی گزینه های انحرافی و بی تاثیر (یعنی: سئوالاتی که همه به آنها پاسخ داده و یا بدون پاسخ بوده اند) ، واریانس ، انحراف معیار، میانگین، میانه نمرات ، اعتبار آزمون ، کمترین نمره، بیشترین نمره، خطای معیار اندازه گیری، ضریب همبستگی کودر – ریچاردسون.

واژه های کلیدی که آموزش داده خواهند شد:

اندازه گیری، آزمون، ارزشیابی (تشخیصی، تکوینی و پایانی) ویژگی های: عینیت، اعتبار، روایی و قدرت تمیز در سوالات آزمون، مشخصات و کاربرد هر یک از انواع آزمونها، روایی ظاهری و محتوایی سوالات آزمون، تحلیل سوالات آزمون / مراحل تحلیل سوال / محاسبه ضریب دشواری سوال / تفسیر ضریب دشواری / محاسبه ضریب تمیز سوال / تفسیر ضریب تمیز سوال / همبستگی میان سوال های آزمون / رابطه بین ضریب دشواری و ضریب تمیز سوال / تجدید نظر و اصلاح آزمون / منحنی ویژگی سوال و استفاده از آن برای تحلیل سوال.

برای تشریح و تجزیه و تحلیل نتایج آزمون، لازم است فراگیران مختصری با انواع آزمون، انواع سوالات (مناسب و نامناسب)، چک لیست و ارسای سوالات طراحی شده و بودجه بندی سوالات آزمون آشنا شوند.

آشنایی با تعاریف و اصطلاحات مورد نیاز در تجزیه و تحلیل نتایج آزمون:

آزمون: آزمون ابزاری است برای به دست آوردن نمونه ای از رفتار فرد. هر آزمون سودمند، ویژگی یار رفتار مورد نظر را با دقت اندازه گیری می کند.

انواع سوال: سوالات یک آزمون می توانند به صورت پاسخ آزاد (انشایی) و یا چندگزینه ای (انتخاب پاسخ از بین دو یا چند پاسخ) باشد.

تحلیل سوال: با تحلیل سوال، هر یک از سوالات به طور جداگانه ارزشیابی، (روایی validity) و پایایی (ثبات reliability) آزمون مشخص می شود.

اندازه گیری: اندازه گیری عبارت است از اختصاص اعداد باروشی بسیار دقیق و قابل تکرار به افراد به شیوه ای منظم، برای نشان دادن ویژگیهای آنها.

روایی validity: آزمونی روایی دارد که برای اندازه گیری آنچه مورد نظر است کافی و مناسب باشد. روایی شرط کافی برای پایایی است.

پایایی یا ثبات reliability: پایایی یعنی ثبات در اندازه گیری های متعدد. برای تعیین ثبات، حداقل باید نتایج دو اندازه گیری را با یکدیگر مقایسه کرد.

ضریب اعتبار: نشان دهنده این است که آزمون تا چه اندازه در دستیابی به هدفهای معین موفق بوده است.

واریانس: واریانس شاخص پراکندگی نمره ها می باشد و میزان تفاوت نمره ها از یکدیگر را توصیف می کند.

واریانس نظامدار: عبارت است از تغییر پذیری اندازه ها به سبب برخی تاثیرهای معلوم یا نامعلوم که «موجب می شود» نمره ها به یک جهت بیش از جهت دیگر گرایش پیدا کنند.

واریانس خطا: واریانس خطا نوسان یا تغییر پذیری اندازه ها بر اثر عوامل تصادفی است. واریانس خطا واریانس تصادفی است. آن تغییر پذیری شاخصهایی است که معمولاً از نوسانهای کوچک و خود جبران کنند (self-compensating) اندازه ها - اینجا، آنجا، بالا، پایین - ناشی می شود.

مواردی که احتمالاً در واریانس نمره های آزمون موثرند:

- ویژگیهای پایدار فرد (مانند: مهارت های عمومی، توانایی درک دستورالعمل، توانایی حل مسایل کلی، اعتماد به نفس، سلامتی، امکانات و تسهیلات و...)

ضریب پایایی (ثبات) stability: همبستگی بین نمره هایی است که در اثر دوبار اجرای یک آزمون حاصل می شود و نشان می دهد که چند درصد از آوریانس آزمون، ناشی از آوریانس خطا نیست. ضریب پایایی با افزایش طول آزمون (یعنی با افزایش تعداد سوالات آزمون)، افزایش می یابد. این رابطه چندان در مورد روایی صدق نمی کند. **ضریب پایایی با افزایش آوریانس نیز افزایش می یابد.** «پایایی» شرط لازم «روایی» است.

عینیت: چنانچه نظر شخصی افراد در تصحیح سوالات هیچگونه دخالتی نداشته باشد گویند آزمون عینی است یا عینیت دارد.

ضریب تمییز: اگر آزمون شونده ای از گروه قوی (یعنی ۲۷ درصد افرادی که بالاترین نمره را گرفته اند) به سوال معینی پاسخ درست دهد و آزمون شونده ای از گروه ضعیف (یعنی ۲۷ درصد افرادی که پایین ترین نمره را گرفته اند) به همین سوال پاسخ نادرست داده باشد سوال مورد بحث دارای قدرت تشخیص آزمون شونده قوی از ضعیف است و موجب بالابردن پایایی و روایی آزمون می شود.

تجزیه و تحلیل نتایج آزمون جهت بررسی روایی و اعتبار آزمون

روشهای تحلیل سوال

الف) تحلیل سوال با استفاده از شاخصهای دشواری و قدرت تمییز:

شاخصهای قدرت تمییز و همبستگی دورشته ای نقطه ای بین مجموع نمره های آزمون و هریک از سوالها، درجه همبستگی یک سوال را با سوالهای دیگر آزمون نشان می دهد. این شاخص نشان می دهد آیا شخصی که در کل آزمون عملکرد خوبی دارد (فردی که ظاهراً نمره بالایی در صفت مورد اندازه گیری دارد) در مقایسه با فردی که نمره او در کل ضعیف است، به سوال مورد بحث پاسخ صحیح داده است یا خیر؟

به عبارت دیگر، شاخصهای قدرت تمییز و همبستگی دورشته ای بین نمره کل آزمون و سوال، نشان می دهند که آیا سوال، گروهی را که نمره کل آزمون آنها خوب است از گروه دیگری که نمره کل آزمون آنها ضعیف است، تمییز داده است یا خیر؟

آزمون مناسب آزموننی است که بتواند آزموندنیهای با تواناییهای متفاوت را از یکدیگر متمایز سازد.

ضریب دشواری (یا درجه سهولت): نسبت آزموندنیهایی که به یک سوال پاسخ درست داده اند به کل آزموندنیها را درجه دشواری همان سوال گویند. به عبارت دیگر درجه دشواری سوال برای سوال I ، (p_i) نسبت آزموندنیهایی که به سوال پاسخ صحیح داده اند تقسیم بر تعداد کل آزموندنیها، تعریف شده است. و هر قدر این نسبت کوچکتر باشد سوال مشکل تر است و به عکس. مثال: سوالی با درجه دشواری $0/3$ خیلی دشوارتر از سوالی است که درجه دشواری آن $0/8$ می باشد.

(۲) تقسیم بر احتمال پاسخ درست $(1+)$ = سطح دشواری مطلوب

$0/25$ = یک چهارم = احتمال پاسخ درست به سوال ۴ گزینه ای

مثال ۱: سطح دشواری سوال مطلوب در سوالات ۴ گزینه ای برابر است با نصف $(1+0/25)$ یعنی $0/625$

مثال ۲: سطح دشواری سوال مطلوب در سوالات ۲ گزینه ای برابر است با نصف $(1+0/5)$ یعنی $0/75$

مثال ۳: سطح دشواری سوال مطلوب در سوالات ۳ گزینه ای برابر است با نصف $(1+0/33)$ یعنی $0/665$

بهترین سطح دشواری بین $0/3$ تا $0/7$ است زیرا احداً کثرت اطلاع را درباره تفاوت بین آزمون شوندهگان به دست می دهند.

درجه دشواری سوال (p_i)، برای تجزیه و تحلیل سوال بسیار مفید است، چنانچه p_i نزدیک به صفر یا یک باشد، سوال باید بطور کلی حذف شود. زیرا هیچ نوع اطلاعاتی در مورد تفاوت‌های آزمودنی‌ها در زمینه توانایی‌ها یا اختلاف سطوح رفتارهای مورد اندازه گیری به ما نمی دهد. اگر $p_i = 0$ ، هیچیک از آزمودنی‌ها به سوال پاسخ صحیح نداده اند و سوال خیلی دشوار است و واریانس صفر است ($p = 0 * (1 - 0) = 0$) و اصولاً کاربرد ندارد. چنانچه $p_i = 1$ ، همه آزمودنی‌ها به سوال پاسخ درست داده اند و هیچ نوع اطلاعاتی در مورد تفاوت‌های آزمودنی‌ها به ما نمی دهد و واریانس سوال صفر است ($p = 1 * (1 - 0) = 0$)

اگر $p_i = 0.5$ ، سوال حداکثر اطلاعات را درباره تفاوت‌های آزمودنی‌ها به ما می دهد. در این صورت واریانس نمره های سوال یعنی: $0.25 = (1 - 0.5) * 0.5 = p_i (1 - p_i)$ به حداکثر می رسد.

نکته مهم: چنانچه سوالها کاملاً همبسته و دارای درجات دشواری 0.5 باشند، در این صورت لازم است نمره کل نیمی از آزمودنی‌ها صفر و نیمی دیگر کامل باشد. در این حالت تفاوت بین سطوح مختلف «رفتار مورد اندازه گیری» قابل تشخیص نخواهد بود.

در آزمونهای صحیح-غلط ضریب دشواری $p = 0.5$ زمانی به دست می آید که همه آزمون شوندگان با حدس زدن به سوال جواب بدهند.

بنابراین بهترین شیوه، انتخاب سوالهایی است که میانگین دامنه تغییر دشواریهای آنها 0.5 باشد. این دامنه برای آزمونهای نامتجانس محدود تر و برای آزمونهای متجانس گسترده تر است.

شاخص تمیز سوال d_i ، شاخص تمیز سوال (d_i) برای سوال I با استفاده از فرمول زیر محاسبه می شود:

$$d_i = \frac{U_i}{N_{iu}} - \frac{L_i}{N_{il}}$$

U_i = تعداد آزمودنی‌هایی که به سوال I پاسخ صحیح داده اند و نمره کل آنها در دامنه بالای نمره های کل آزمون قرار می گیرد.

L_i = تعداد آزمودنی‌هایی که به سوال I پاسخ صحیح داده اند و نمره کل آنها در دامنه پایین نمره های کل آزمون قرار می گیرد.

N_{iu} = تعداد آزمودنی‌هایی است که نمره کل آنها در دامنه بالای نمره های کل آزمون قرار دارد.

N_{il} = تعداد آزمودنی‌هایی است که نمره کل آنها در دامنه پایین نمره های کل آزمون قرار دارد.

در صورتی که: $N_{iu} = N_{il} = N_i$ ، فرمول بالا بصورت زیر کاهش می یابد:

$$d_i = \frac{U_i - L_i}{N_i}$$

دامنه تغییر ضریب تمیز از +1 تا -1 می باشد، که اعداد مثبت (+) مطلوب ما است. اعداد منفی (-) نشان می

دهد آزمون شونده قوی تر کمتر از آزمون شونده ضعیف تر به سوال مزبور پاسخ درست داده اند یعنی سوال قدرت تمیز ندارد و گمراه کننده یا مبهم است.

هنگامی که برای سوالهای چندگزینه ای تحلیل سوال انجام می شود، بررسی شاخصهای دشواری و تمیز یا همبستگی سوال با نمره کل آزمون، معمولاً با ارزش هستند. به همین ترتیب نسبت آزمودنی‌هایی که

در گروه‌های مختلف توانایی قراردارندوهریک از گزینه‌ها را اعلامت زده اند، ارزشمنداست (پاسخهای انتخابی نادرست گزینه‌های انحرافی نامیده می‌شوند).

در جدول شماره (۱) آزمودنیها براساس نمره‌های کل آزمون به سه نمونه بالا (U)، متوسط (M) و پایین (L) طبقه بندی شده‌اند.

گزینه‌هایی که با ستاره علامتگذاری شده، گزینه صحیح است و سه پاسخ دیگر گزینه‌های انحرافی هستند. سوال یک با $P_i=0/37$ (درجه دشواری)، $d_i=0/5$ (شاخص تمیز) و $r_{pbis}=0/45$ (همبستگی دورشته‌ای) سوال نسبتاً خوبی است. سه گزینه انحرافی این سوال (d و c و a) خوب عمل کرده‌اند.

جدول شماره (۱)

سوال سه			سوال دو			سوال یک					
گزینه	M	L	U	گزینه	M	L	U	گزینه	M	L	U
a	0/07	0/03	0/00	*a	0/70	0/30	0/10	a	0/50	0/30	0/20
b	0/04	0/04	0/00	b	0/30	0/70	0/90	*b	0/10	0/40	0/60
*c	0/80	0/90	1/00	c	0/00	0/00	0/00	c	0/20	0/15	0/10
d	0/09	0/03	0/00	d	0/00	0/00	0/00	d	0/20	0/15	0/10
$P_i=0/90$ $d_i=0/20$ $r_{pbis}=0/23$			$P_i=0/37$ $d_i=0/60$ $r_{pbis}=0/53$			$P_i=0/37$ $d_i=0/50$ $r_{pbis}=0/45$					

* پاسخ صحیح

گزینه‌های انحرافی هرپرسش چندگزینه‌ای در صورتی دارای کارایی است که دویژگی زیرادار باشد:

- توجه افرادی اطلاع از موضوع اندازه‌گیری را به اندازه گزینه درست به خود جلب کنند.
- برای آزمودنیهایی که در مورد موضوع اندازه‌گیری توانایی و مهارت کافی دارند گول زنده نباشند. علاوه بر آن، افراد گروه بالا باید بیش از افراد گروه پایین گزینه درست را انتخاب کنند، و این تفاوت باید از نظر آماری معنی‌دار باشد.

در جدول شماره ۲ تعداد افراد گروه بالا و پایین که به ۵ سوال پاسخ داده‌اند و در آن گزینه درست با علامت * مشخص شده آمده است:

جدول شماره (۲)

شماره سوال	گروه بالا						گروه پایین						تعداد پاسخ‌های درست	سطح دشواری سوال
	الف	ب	ج	د	بدون جواب	جمع	الف	ب	ج	د	بدون جواب	جمع		
۱	۰	۲۵*	۰	۰	۰	۲۵	۲	۱۸*	۱	۲	۲	۲۵	۴۳	۰/۸۶
۲	۱۴*	۸	۲	۰	۱	۲۵	۰*	۱۵	۵	۳	۲	۲۵	۱۴	۰/۲۸
۳	۴	۱۵	۲*	۴	۰	۲۵	۷	۳	۸*	۶	۱	۲۵	۱۰	۰/۲۰
۴	۰	۲۰*	۰	۵	۰	۲۵	۰	۱۵*	۰	۹	۱	۲۵	۳۵	۰/۷۰
۵	۱	۰	۰	۲۳*	۱	۲۵	۸	۷	۶	۴*	-	۲۵	۲۷	۰/۵۴

سوال شماره ۱: به این سوال ۸۶ درصد افراد دو گروه پاسخ درست داده اند، بنابراین سوال آسانی است. گزینه انحرافی الف برای هیچ یک از افراد گروه بالا جاذبه ای نداشته است اما توجه ۸ درصد افراد گروه پایین را به خود جلب کرده است $\{0.08 = 0.25\}$ تقسیم بر ۲). چون بین درصد پاسخ های دو گروه بالا و پایین به این گزینه از نظر آماری تفاوت معنی دار وجود دارد.

برای آزمون معنی دار بودن نسبت فراوانی پاسخ های دو گروه بالا و پایین به هر گزینه از آزمون نسبت استفاده می شود. **آزمون نسبت:**

$\{Pm(1-Pm)/N\}$ تقسیم بر $T=(P1.P2)$ در این فرمول $P1$ و $P2$ نسبت انتخاب پاسخ انحرافی توسط افراد دو گروه، Pm متوسط نسبت انتخاب این گزینه توسط افراد دو گروه و N مجموع افراد دو گروه است. بنابراین خواهیم داشت: $P1=0, P2=0/08, Pm=(0+2)/50=0/04$ و لذا $t=2/887$ و چون t محاسبه شده از t مبین در آزمون یک دامنه ای (که از روی جدول اعداد استخراج می شود) بزرگتر است. لذا تفاوت بین دو نسبت از نظر آماری معنی دار است.

بنابراین می توان گفت که این سوال با وجود آسان بودن دارای گزینه های انحرافی مناسب و موثر است. **سوال شماره ۲:** سطح دشواری این سوال ۰/۲۸. و لذا سوال دشواری است. همه گزینه های انحرافی توجه افراد را به خود جلب کرده و لذا گزینه های موثری بوده اند. در همه موارد فراوانی و درصد انتخاب گزینه های انحرافی توسط افراد ضعیف بیش از فراوانی و درصد افراد گروه قوی بود و تفاوت بین نسبت انتخابها معنی دار است. بنابراین گزینه های انحرافی از کارایی لازم برخوردار بوده اند.

سوال شماره ۳: این سوال با سطح دشواری ۰/۲۰، سوال دشواری است. چون درصد پاسخ های درست افراد گروه ضعیف بیش از درصد پاسخ های درست افراد گروه قوی است، لذا سوال ضریب تشخیص مناسبی ندارد. علت این امر احتمالاً از آنجاست که می شود که گزینه انحرافی (ب) برای افراد گروه قوی جاذبه زیادی داشته است. به طوری که ۶۰ درصد افراد این گروه $(0.60 = 0.25 / 15)$ آن را به عنوان گزینه درست انتخاب کرده اند، بنابراین گزینه (ب) گول زننده است. با وجودی که گزینه های انحرافی (الف) و (د) از کارایی کافی برخوردار هستند، اما چون سوال فاقد قوه تمیز است یا باید در آن تجدید نظر کلی به عمل آید و یا از آزمون حذف شود.

سوال شماره ۴: سطح دشواری این سوال ۰/۷۰ است که از این نظر با وجود آسان بودن، سوال نسبتاً مناسبی است و گزینه انحرافی (د) از کارایی لازم برخوردار است. اما هیچ یک از گزینه های (الف) و (ج) کارایی لازم را نداشته است. زیرا توجه هیچ یک از آزمون شوندگان، حتی آزمون شوندگان گروه ضعیف را نیز به خود جلب نکرده اند. چنانچه لازم باشد این سوال در آزمون گنجانده شود باید در گزینه های (الف) و (ج) تجدید نظر شود تا به قدر کافی توجه افراد گروه ضعیف را به خود جلب کنند.

سوال شماره ۵: سوال شماره ۵ با داشتن سطح دشواری ۰/۵۴، سوال مناسبی است. به ویژه همه گزینه های انحرافی از کارایی لازم برخوردار بوده اند. زیرا این گزینه های بیشتر توجه افراد گروه پایین را به خود جلب کرده اند. این سوال را به همین صورت می توان در آزمون فرم اصلی به کار بست.

ب) تحلیل سوال با استفاده از شاخص همبستگی دورشته ای نقطه ای:

یکی دیگر از روشهای بررسی ضریب تشخیص سوالها، محاسبه همبستگی بین عملکرد آزمون شوندگان در هر سوال با عملکرد آنان در کل آزمون است. در این همبستگی یکی از متغیرها (عملکرد آزمون شوندگان در هر سوال) دوارزشی و متغیر دیگر (عملکرد آزمون شوندگان در کل آزمون) متغیر پیوسته است.

این نوع همبستگی، در اصطلاح همبستگی دورشته ای نقطه ای نامیده می شود. همبستگی دورشته ای نقطه ای بین سوال و نمره کل آزمون از فرمول زیر محاسبه می شود:

آزمون (r_{ix}): شاخص دیگری که به جای **d_i** برای تمیز سوال بکار برده می شود، همبستگی بین نمره های سوال **i** و نمره کل آزمون **x** با استفاده از فرمول زیر محاسبه می شود:

$$r_{pbis} = \frac{M_p - M_t}{S_t} \cdot \sqrt{\frac{P_i}{1 - P_i}}$$

در این فرمول **r_{pbis}** ضریب همبستگی دورشته ای نقطه ای، **M_p** میانگین نمره های خام آزمون شوندگانی است که به سوال مورد نظر پاسخ درست داده اند، **M_t** میانگین نمره های خام همه آزمون شوندگان، **S_t** انحراف معیار نمره های خام همه آزمون شوندگان، **P_i** نسبت کسانی است که به سوال مورد نظر پاسخ درست داده اند به کل آزمون شوندگان.

مثال ۱: فرض کنید میانگین و انحراف معیار نمره های خام یک آزمون ۷۰ سوالی که در مورد یک گروه نمونه ۵۰ نفری اجرا شده است به ترتیب برابر ۹/۷ و ۵۴/۳ و میانگین نمره های ۲۹ نفر که به سوال شماره ۵ درست پاسخ داده اند برابر ۵۷/۶ باشد. در این صورت خواهیم داشت:

$$p_5 = \frac{29}{50} = 0/58 \text{ و } S_t = 9/7, M_t = 54/3, M_p = 57/6 \text{ ولذا}$$

$$r_{pbis} = \frac{57/6 - 54/3}{9/7} \cdot \sqrt{\frac{0/58}{1 - 0/58}} = 0/40$$

ضریب تشخیص سوال شماره ۵ برابر ۰/۴۰ است. به عبارت دیگر بین موفقیت در سوال شماره ۵ و نمره کل آزمون ۰/۴۰ همبستگی وجود دارد.

مثال ۲: اگر متوسط همبستگی دورشته ای (یعنی: **r_{pbis}**) بین نمره های سوال و نمره کل آزمون بین ۰/۳ تا ۰/۴ باشد، دامنه ای بین **p_i = ۰/۴** تا **p_i = ۰/۶** را پیشنهاد می کند (هنریسن صفحه ۱۹۷۱، ۱۵۳).

در صورتی که همبستگی دورشته ای سوال و کل آزمون، بالاتر از ۰/۴ باشد (اگر آزمون دارای تجانس بیشتری باشد) درجات دشواری با دامنه گسترده تری پیشنهاد می شوند.

توجه: برای آزمون معنی دار بودن این همبستگی دوروش وجود دارد:

۱- اگر تعداد آزمون شوندگان زیاد باشد (۳۰ نفر و بیشتر) ابتدا خطای معیار ضریب همبستگی دورشته ای نقطه ای را از فرمول:

$$S_{r_{pbis}} = \frac{1}{\sqrt{N}}$$

محاسبه می کنیم. برای آزمون معنی دار بودن ضریب محاسبه شده در یکی از دو سطح ۰/۰۵ یا ۰/۰۱، خطای معیار ضریب همبستگی دورشته ای نقطه ای را در اندازه های ۰/۰۵ Z (۱/۹۶) و یا ۰/۰۱ Z (۲/۵۸) ضرب می کنیم

تفاوت ضریب همبستگی دورشته ای محاسبه شده از صفر در سطح ۰/۰۵ یا ۰/۰۱ به دست آید. ضریب همبستگی محاسبه شده در صورتی معتبر خواهد بود که در سطح ۰/۰۱ و یا ۰/۰۵ معنی دار باشد. به عبارت دیگر اگر داشته باشیم:

$$1/96 * Sr_{pbis} > r_{pbis} > 2/58 * Sr_{pbis}$$

در این صورت ضریب همبستگی محاسبه شده در سطح ۰/۰۵ معنی دار است. هرگاه داشته باشیم: $r_{pbis} > 2/58 * Sr_{pbis}$ ، در این صورت ضریب همبستگی محاسبه شده در سطح ۰/۰۱ معنی دار است. اگر ضریب محاسبه شده از $1/96 * Sr_{pbis}$ کوچکتر و یا برابر آن باشد، ضریب همبستگی محاسبه شده در هیچ سطحی معنی دار نیست.

۲- اگر تعداد آزمون شوندگان کم باشد ($N < 30$) برای آزمون معنی دار بودن ضریب همبستگی دورشته ای نقطه ای محاسبه شده برای هر سوال، تفاوت $Mp - Mt$ را از راه آزمون t بررسی می کنیم.

تذکر مهم: در آزمونهایی که تعداد سوالهای آن کم است، استفاده از همبستگی دورشته ای نقطه ای برای تعیین ضریب تشخیص سوالها با اشکال مواجه خواهد بود، زیرا موفقیت در سوالی که قرار است ضریب همبستگی آن بانمره کل آزمون محاسبه شود نمره آزمون شوندگان را در کل آزمون افزایش می دهد.

مثال: در آزمونی که ۶ سوال دارد، پاسخهای درست آزمون شوندگان به سوالی که قرار است همبستگی آن بانمره کل آزمون محاسبه شود، موجب افزایش نمره کل آزمون می شود. در نتیجه ضریب همبستگی محاسبه شده افزایش می یابد. برای اجتناب از این اشکال لازم است که نمره هر سوال از نمره کل آزمون حذف شود (کاپلان و ساکوزو، ۱۹۸۹) در جدول شماره (۳) مراحل انجام این عمل نشان داده شده است:

در این جدول پاسخهای یک گروه ۱۰ نفری از آزمون شوندگان به یک آزمون ۶ سوالی را نشان می دهد. در ستون افقی جدول، شماره افراد در ستون عمودی شماره سوالها و در خانه های جدول پاسخ هر آزمودنی به هریک از سوالها بانمره های ۱ و ۰ نشان داده شده است. نمره ۱ نشان می دهد آزمون شونده به سوال مورد نظر پاسخ درست داده است و نمره صفر (۰)، نشانگر آن است که آزمون شونده به سوال پاسخ غلط داده است. در ستون جمع، تعداد پاسخهای درست به هر سوال نوشته شده است، و در سطر آخر جدول مجموع نمره های خام هر آزمون شونده در کل آزمون (مجموع نمره های نوشته شده است).

جدول شماره (۳)

شماره افراد شماره سوال	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	جمع
۱	۱	۱	۰	۱	۱	۱	۰	۱	۱	۱	۸
۲	۱	۱	۱	۱	۱	۱	۰	۱	۱	۱	۹
۳	۱	۱	۱	۰	۱	۱	۰	۱	۱	۱	۸
۴	۱	۰	۱	۱	۰	۱	۰	۰	۰	۱	۵
۵	۱	۱	۱	۱	۱	۱	۰	۰	۰	۱	۷
۶	۱	۱	۰	۱	۱	۱	۰	۰	۰	۰	۵
نمره خام	۶	۵	۴	۵	۵	۶	۰	۳	۳	۵	۴۲

محاسبه ضریب همبستگی دورشته ای نقطه ای برای سوال شماره ۱:

باتوجه به توضیح مندرج در مثال بالای جدول و با در نظر گرفتن جدول شماره ۳ نمره های آزمون شوندگان شماره ۱ تا ۱۰ در کل آزمون با حذف نمره سوال شماره ۱ به ترتیب برابر ۵، ۴، ۴، ۴، ۴، ۵، ۴، ۲، ۲، ۰ است که میانگین

آنها (M_t) برابر $3/4$ و انحراف معیار آنها (St) برابر $1/58$ است. نسبت پاسخهای درست آزمون شوندگان به سوال شماره ۱ (P_1) برابر $0/8$ است. نمره های خام آزمون شوندگانی که به سوال ۱ درست پاسخ داده اند با حذف نمره سوال شماره ۱ به ترتیب برابر $4-2-2-5-4-4-4-5$ و میانگین آنها (M_p) برابر $3/75$ و انحراف معیار آنها برابر $1/16$ است. بنابراین خواهیم داشت:

$$r_{pbis} = \frac{3/75 - 3.4}{1/58} \cdot \sqrt{\frac{.8}{(1 - .8)}} = .44$$

برای آزمون معنی دار بودن ضریب همبستگی محاسبه شده تفاوت $M_p - M_t$ را از فرمول زیر بررسی می کنیم:

$$t = \frac{M_p - M_t}{\sqrt{\frac{S^2_p}{N_p} + \frac{S^2_t}{N_t} - 2r_{pbis} \left(\frac{S_p}{N_p}\right) \left(\frac{S_t}{N_t}\right)}}$$

با جایگذاری مقادیر عددی در فرمول بالا خواهیم داشت: $t = 0/718$

چون t محاسبه شده از t مبین (که از جدول استخراج می شود) در سطح $0/05$ (یعنی از $1/746$) کمتر است، لذا فرض همبستگی صفر (0) بین سوال شماره ۱ و نمره کل آزمون رانمی توان رد کرد. بنابراین ضریب همبستگی محاسبه شده معنی دار نیست و این سوال با سایر سوالات آزمون همسانی درونی ندارد و باید از آزمون حذف شود. محاسبه ضریب همبستگی دورشته ای نقطه ای اگر تعداد سوالها زیاد باشد بدون استفاده از رایانه کاری وقت گیر و پر زحمت است.

به طور کلی ضریب همبستگی دورشته ای نقطه ای، همبستگی بین یک متغیر دوارزشی (0 و 1) و یک متغیر پیوسته رانشان می دهد. برخلاف کاربرد ضریب همبستگی گشتاوری پیرسون (که شرط استفاده از این همبستگی آن است که توزیع نمره ها باید در هر دو متغیر نرمال باشد)، در کاربرد ضریب همبستگی دورشته ای نقطه ای شرط نرمال بودن توزیع متغیر دوارزشی لازم نیست.

مثال: اگر بخواهیم میزان همبستگی بین جنسیت و قد را محاسبه کنیم، می توانیم بین این دو متغیر همبستگی دورشته ای نقطه ای را بکار ببندیم. در این مثال، قد یک متغیر پیوسته است، اما جنسیت (زن یا مرد بودن) یک متغیر دوارزشی است که توزیع این صفت در جامعه نرمال نیست.

نکته دیگری که در استفاده از همبستگی دورشته ای نقطه ای باید به خاطر داشت این است که اگر M_p, M_t, St از روی نمره های ملاک خارجی (مثلاً "نمره های امتحانهای رسمی) محاسبه شده باشد، در این صورت همبستگی محاسبه شده شاخص **روایی (validity)** ملاکی سوال است. هرگاه اندازه های مذکور از روی نمره های آزمون محاسبه شود، همبستگی به دست آمده شاخص همسانی درونی سوال با سایر سوالها و کل آزمون است. بدیهی است چنانچه آزمون دارای **روایی** باشد، در این صورت همبستگی محاسبه شده شاخصی از **روایی آزمون** نیز خواهد بود.

روایی سوالهای آزمون:

ضریب همبستگی دورشته ای نقطه ای به عنوان شاخصی از روایی یک سوال برای پیش بینی احتمال موفقیت آزمون شونده در یک ملاک خارجی رابه صورت زیرنیزمی توان مطرح کرد(ایکن، ۱۹۸۵)

$$r_{pb} = \frac{\bar{Y}_1 - \bar{Y}}{s} \cdot \sqrt{\frac{n_1 n}{(n - n_1)(n - 1)}}$$

در این فرمول، n تعداد کل آزمون شوندگان، n_1 تعداد آزمون شوندگانی است که به سوال مورد نظر درست پاسخ داده اند. \bar{Y} میانگین نمره های خام همه آزمون شوندگان در متغیر ملاک، و \bar{Y}_1 میانگین نمره های خام متغیر ملاک در مورد آزمون شوندگانی است که به سوال مورد نظر پاسخ درست داده اند، s انحراف معیار نمره های خام همه آزمون شوندگان در متغیر ملاک است. البته فرمول را برای تعیین همبستگی سوال با کل آزمون نیز می توان به کار برد. در این صورت متغیر ملاک نمره های کل آزمون برای آزمون شوندگان خواهد بود.

مثال:

- ۱- فرض کنید آزمونی برای پیش بینی احتمال موفقیت در دروسهای کتبی امتحانهای نهایی دبیرستان ساخته شده است. در این صورت متغیر ملاک، معدل امتحانهای کتبی گروه نمونه ای از آزمون شوندگانی است که آزمون مورد بحث درباره آنها اجرا شده است.
- ۲- فرض کنید از یک گروه نمونه ۳۰۰ نفری که در این آزمون شرکت کرده اند، ۱۷۰ نفر به سوال شماره ۵ درست پاسخ داده باشند. هرگاه میانگین معدل امتحانهای کتبی این ۱۷۰ نفر ۱۶، و میانگین و انحراف معیار معدل امتحان کتبی نهایی برای همه ۳۰۰ نفر به ترتیب برابر ۱۲ و ۵ باشد، همبستگی دورشته ای نقطه ای سوال ۵ به شرح زیر محاسبه می شود:

$$r_{pb} = \frac{16 - 12}{5} \cdot \sqrt{\frac{170 (300)}{(300 - 170)(300 - 1)}} = 0.92$$

- ۱- برای آزمون معنی دار بودن این ضریب، می توان از روشی که پیش از این در مورد همبستگی دورشته ای نقطه ای توضیح داده شد استفاده کرد.
- ۲- هر قدر این همبستگی به ۱ نزدیکتر باشد، قدرت پیش بینی عملکرد آزمون شوندگان در متغیر ملاک به وسیله سوال مورد نظر بیشتر خواهد بود.
- ۳- سوالهایی که دارای شاخص روایی پیش بینی بالاتری هستند، هر چند همبستگی آنها با سایر سوالهای آزمون کم باشد، برای گنجاندن در آزمون فرم نهایی مناسب تر خواهد بود.

ج) تحلیل سوال با استفاده از شاخص ضریب همبستگی دورشته ای:

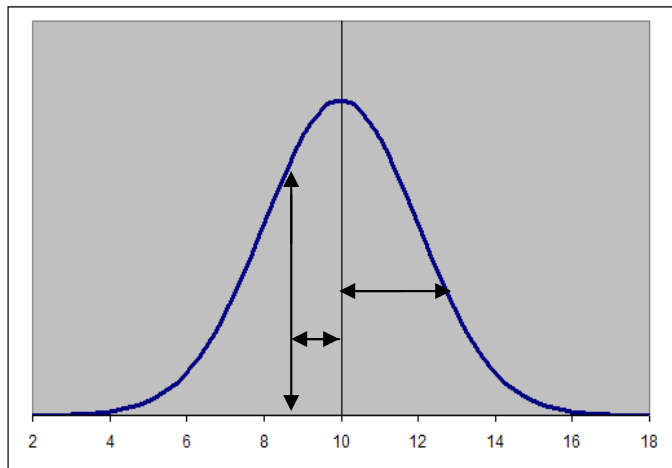
ضریب همبستگی دورشته ای نیز مانند ضریب همبستگی دورشته ای نقطه ای برای محاسبه همبستگی یک متغیر پیوسته مثلاً "نمره های آزمون با یک متغیر دوارزشی یعنی هریک از سوالهای آزمون به کار می رود. برخلاف همبستگی دورشته ای نقطه ای، شرط استفاده از ضریب همبستگی دورشته ای آن است که توزیع نمره هادر هر دو متغیر (خواه متغیر پیوسته و خواه متغیر دوارزشی) در جامعه نرمال باشد. ضریب همبستگی دورشته ای از رابطه زیر به دست می آید:

$$rbis = \frac{Mp - Mt}{St} \cdot \frac{p}{Y}$$

در این فرمول، Mp میانگین نمره های آزمون یا نمره های ملاک کسانی است که به سوال مورد نظر پاسخ درست داده اند. St و Mt به ترتیب میانگین و انحراف معیار نمره های آزمون یا نمره های ملاک همه آزمون شوندگان است. p نسبت کسانی است که به سوال پاسخ درست داده اند. Y عرض نقطه تفکیک معادل p در منحنی توزیع نرمال به مساحت یک واحد است.

مثال: فرض کنید از صد (۱۰۰) آزمون شونده ای که در یک آزمون شرکت کرده اند ۴۰ نفر به سوال شماره ۱ درست پاسخ داده باشند. هرگاه میانگین نمره های این ۴۰ نفر در آزمون برابر ۶۵ و میانگین و انحراف معیار همه ۱۰۰ نفر در همان آزمون به ترتیب ۱۲ و ۶۰ باشد، ضریب همبستگی دورشته ای برای سوال شماره ۱ به شرح زیر محاسبه می شود: $p = \frac{40}{100} = 0/40$ و $St=12$ و $Mt=60$ و $Mp=65$

چنانچه



$$rbis = \frac{Mp - Mt}{St} \cdot \frac{p}{Y} = \frac{65 - 60}{12} * \frac{0/40}{0/386} = 0/43$$

د) تحلیل سوال با استفاده از شاخصهای پایایی (ثبات) و روایی سوال:

آزمونی با N سوال رابطه آزمایشی یک بار اجرا کرده ایم و داده های آن را در اختیار داریم و می خواهیم از بین آنها K سوال را به عنوان بهترین سوال انتخاب کنیم. به عبارت دیگر این k سوال، باید پایایی (ثبات) درونی (r_{xy}) یا روایی و وابسته، به حداکثر r_{xy} برسد.

در این بحث X^* برای نشان دادن نمره کل آزمونی که از N سوال X و نمره آزمونی است که از k سوال تشکیل شده است و Y نمره ملاک مورد علاقه است.

برای تشخیص بهترین سوال محاسبه چهار شاخص آماری به شرح زیر برای هر سوال I ، لازم است:

۱. دشواری سوال (p_i)

۲. انحراف معیار سوال $\{ s_i = \sqrt{p_i(1-p_i)} \}$

۳. شاخص پایایی سوال $S_i F_{ix}^*$ (در اینجا، F_{ix}^* همبستگی دورشته ای نقطه ای بین سوال و نمره آزمونی است که از N سوال تشکیل شده است)

۴. شاخص روایی سوال ($S_i F_{iy}$) که در آن F_{iy} همبستگی دورشته ای نقطه ای بین نمره های سوال و ملاک است).

وقتی این چهار شاخص آماری را شناختیم، می توانیم با محاسبه میانگین، انحراف معیار، پایایی (ثبات) و روایی نمره های X آزمون، رابراساس k سوال انتخاب شده، برآورد کنیم: بهترین سوال، سوالی است که شاخصهای روایی و پایایی آن به هم نزدیک باشند. این گونه سوالها درست راست بالای نمودار قرار خواهند گرفت.

شاخصهای مرکزی (متوسطها):

۵. میانگین، از تقسیم مجموع نمره ها بر تعداد نمره ها به دست می آید.
۶. میانه، نمره وسط است پس از سورت نمره ها. (اگر وقت کافی برای محاسبه میانگین نباشد کاربرد دارد).
۷. نما، نمره ای است که بیشتر از همه نمره ها، تکرار شده است. و یک برآورد تقریبی و سریع از ارزش مرکزی است.

شاخصهای پراکندگی:

۱. واریانس $S^2(x) = 1/N \{ \sum x^2 - N(\sum x/N)^2 \}$

{ (میانگین نمره ها) * (تعداد نمره ها) - (مجموع مربعات نمره ها) = واریانس

۲. انحراف معیار $S(x) = \sqrt{1/N \{ \sum x^2 - N(\sum x/N)^2 \}}$

جذر واریانس = انحراف معیار

۳. دامنه تغییر: تفاوت بیشترین و کمترین نمره را گویند. سریعترین شاخص پراکندگی است و درباره کسب اطلاعات مورد نیاز از نمره های انتهایی کاربرد دارد. اگر تعداد نمره ها بیش از ۱۰۰ عدد باشد از تقسیم دامنه تغییر بر ۶ یک انحراف معیار (تقریبی) به دست می آید.

ضریب همبستگی: (r_{xy}) شاخصی از نیرومندی بین دو متغیر است. و رایجترین تفسیر همبستگی براساس

مجذور آن (r_{xy}^2) انجام می شود. و برابر است بانسبت واریانس y از x که به وسیله همبستگی خطی با x برآورد می شود. در ساده ترین شکل خود، عددی است که میزان رابطه بین دو متغیر را تعیین می کند. این شاخص میزان هماهنگی تغییرات یک متغیر با تغییرات متغیر دیگر اندازه می گیرد.

اگر ضریب همبستگی = ۱، یعنی همبستگی دو متغیر کامل و مثبت است و بدون استثنا، تغییرات یک متغیر هم جهت و همسان با تغییرات متغیر دیگر است. و اگر ضریب همبستگی = صفر، یعنی تغییرات یک متغیر ارتباطی با تغییرات متغیر دیگر ندارد.

اگر ضریب همبستگی = -1، یعنی بین دو متغیر همبستگی کامل و منفی وجود دارد و بدون استثنا، تغییرات یک متغیر همسان با تغییرات دیگر و در جهت معکوس است.

ضریب همبستگی مستلزم دو بار اندازه گیری یک گروه آزمودنی در دو متغیر یا گروه های هم‌تاست و نمی توان آن را برای یک متغیر محاسبه کرد.

مثال: نمرات آزمون کتبی ۱۲ نفر از آزمون شوندگان پایان یک دوره آموزشی به شرح زیر است: $X_i = 51-50-52-51-65-50-$

$50-55-90-77-74-80-75-72$

نمرات آزمون عملی آنها نیز به ترتیب به شرح زیر است: $Y_i = 74-70-60-72-70-60-80-80-40-70-70-60$

برای محاسبه ضریب همبستگی بین نمرات کتبی و عملی آزمودنی ها به شرح زیر عمل می شود:

$$R_{xy} = \frac{S_{xy}}{S_x * S_y}$$

X=نمره کتبی	Y=نمره عملی	xy	X ²	Y ²
51	60	3060	2601	3600
52	70	3640	2704	4900
50	70	3500	2500	4900
65	40	2600	4225	1600
72	80	5760	5184	6400
75	80	6000	5625	6400
80	60	4800	6400	3600
74	70	5180	5476	4900
77	72	5544	5929	5184
90	60	5400	8100	3600
55	70	3850	3025	4900
50	74	3700	2500	5476
$\sum X=791$	$\sum Y=806$	$\sum xy=53034$	$\sum X^2=54269$	$\sum Y^2=55460$

$$R_{xy} = \frac{4419,5 - (65,916 * 67,166) - 7/8}{\sqrt{77.516 * 110.305}} = \frac{-7/8}{\sqrt{8550.402}} = -0,084$$

عدد 084/-. نشان دهنده وجود همبستگی معکوس بین نمرات کتبی و عملی آزمون مورد نظرمی باشد.
توجه: تمام محاسبات انجام شده برای آشنایی با شکل محاسبات است و نیازی به حفظ کردن فرمول نمی باشد زیرا انجام همه این موارد توسط رایانه انجام خواهد شد.

ه) تحلیل سوال با استفاده از منحنی ویژگی سوال:

منحنی ویژگی سوال نمایشی ترسیمی از رابطه بین احتمال پاسخگویی به یک سوال معین و وضعیت آزمودنی در رفتار مورد اندازه گیری آزمون است.

چون نمره های رفتار مورد مطالعه موجود نیستند، نمره های مشاهده شده برای برآورد ارزشهای این رفتار به کار برده می شوند.

برای هر سوال «منحنی ویژگی سوال» از طریق ترسیم نمره های کل آزمون بر محور افقی و نسبت آزمودنیهایی که به سوال پاسخ صحیح داده اند، بر روی محور عمودی برآور می شوند.

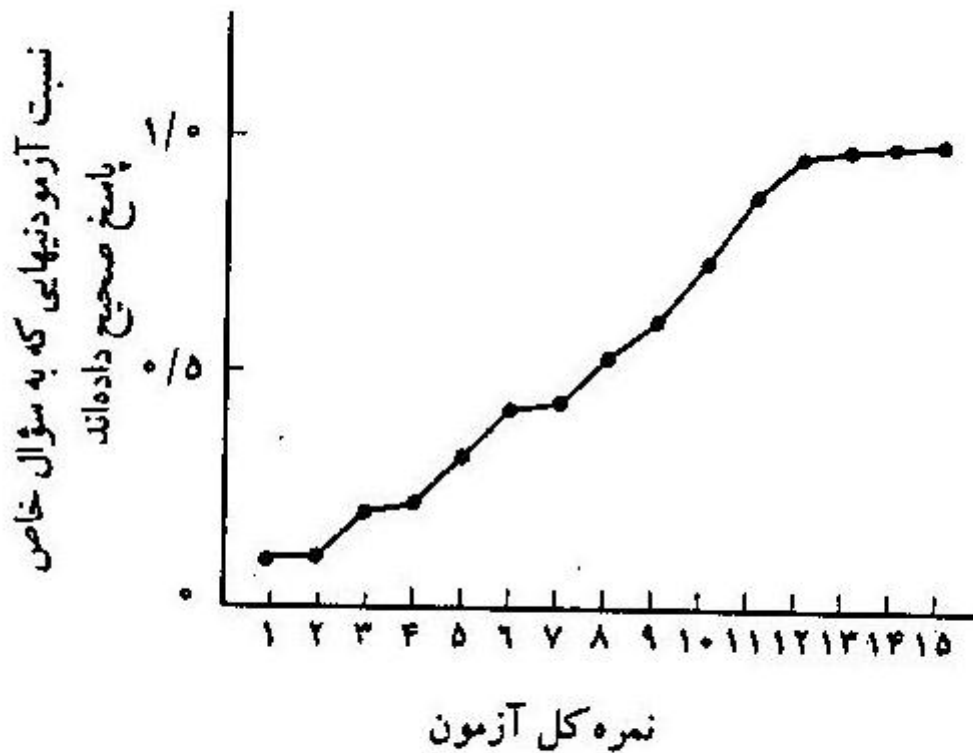
منطق تحلیل سوال با استفاده از منحنیهای ویژگی سوال همانند تحلیل سوال از طریق شاخصهای دشواری و قدرت تمیز سوال است.

منحنی ویژگی سوال، برآورد شده برای پنج سوال در شکل زیر نشان داده شده است:

سوال A، سوال ضعیفی است، زیرا توانایی تمیز بین آزمودنیها با تواناییهای مختلف را ندارد.

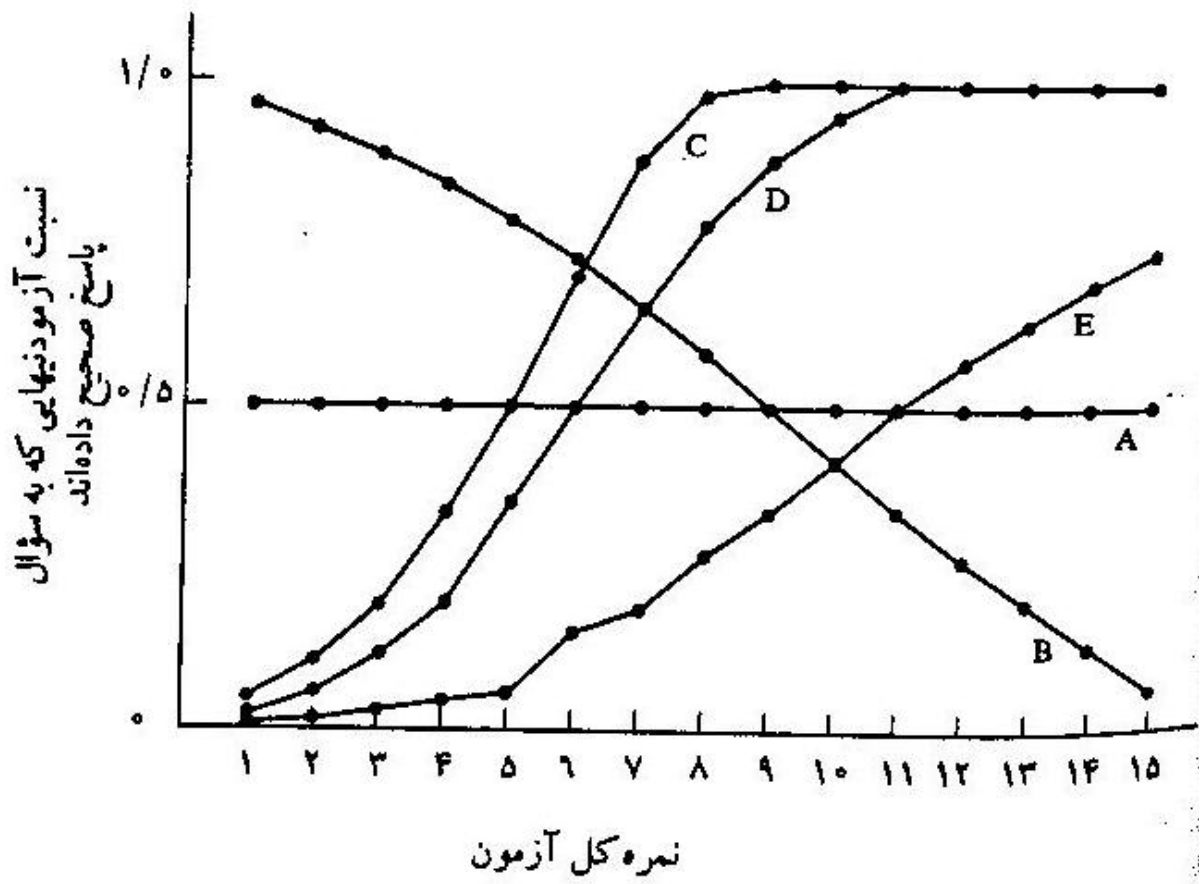
سوال B سوال ضعیفی است، زیرا بین این سوال و نمره کل آزمون رابطه منفی وجود دارد. (منحنی

ویژگی سوال برآورد شده برای این سوال همان انتظاری است که در مورد نسبت اشخاصی داریم که گزینه غلط را انتخاب می کنند.)



شکل ۳-۶ منحنی ویژگی سؤال برآورد شده

سوالهای C و D، سوالهای قابل قبولی هستند. منحنی ویژگی سؤال E مسطح تر از منحنیهای ویژگی سوالهای C و D سوالهای قابل قبولی هستند. منحنی ویژگی سؤال E مسطح تر از منحنیهای ویژگی سوالهای C و D است. این عامل نمی تواند تمیزدهنده کارایی برای آزمودنی‌هایی باشد که نمره آنها اختلاف زیادی ندارد. شیب منحنی ویژگی سؤال می تواند به عنوان یک شاخص تمیزدهنده به کار برده شود، برای مثال منحنی ویژگی سؤال برآورد شده در شکل زیر بین اعداد ۷ و ۸ دارای شیب زیاد است. این سوال بین نمره آزمودنی‌های پایین از ۸ و نمره های مساوی یا بزرگتر از ۸ تمیزدهنده قدرتمندی است.

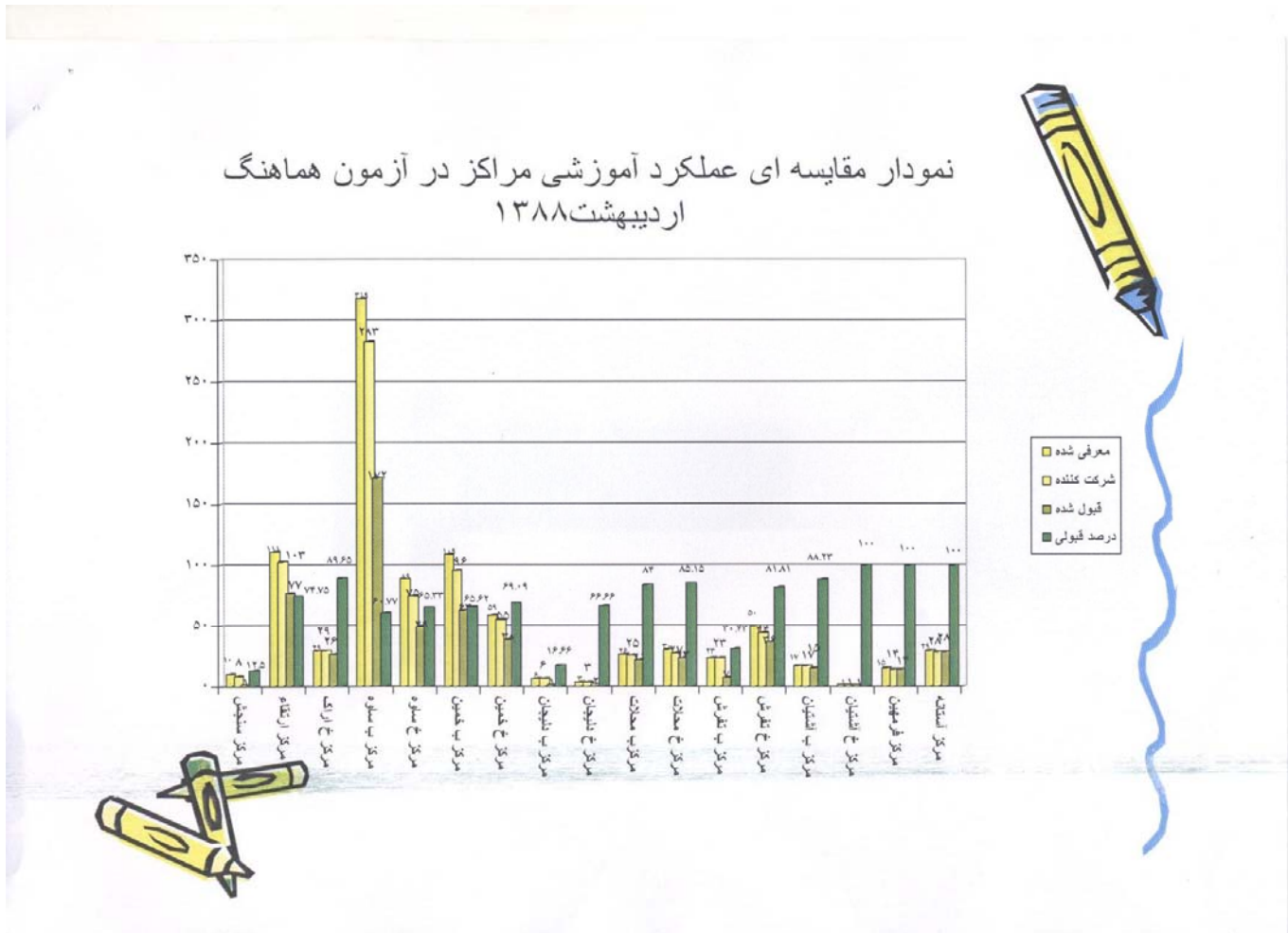


شکل (۴-۶)

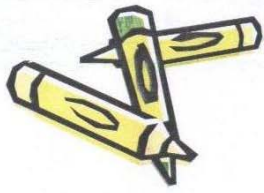
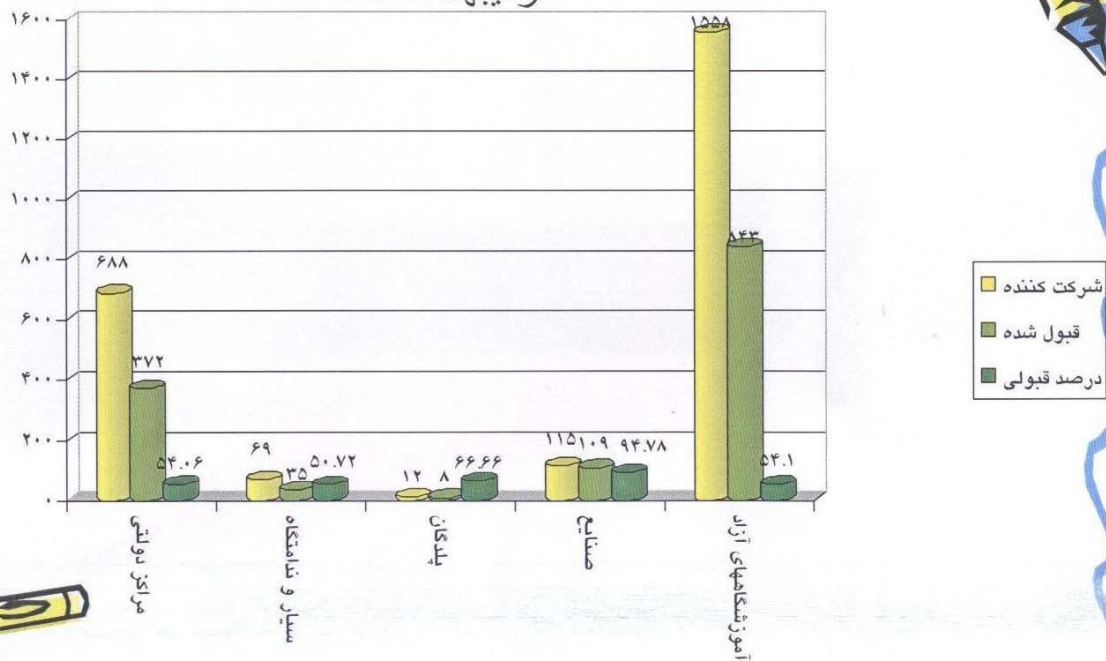
تجزیه و تحلیل نتایج آزمون جهت بررسی تعداد قبول شدگان و مقایسه عملکرد آموزشی:

با استفاده از نرم افزار مدیریت آزمون، می توان نتایج آزمون کتبی را به تفکیک: شهرستان، مرکز، حرفه و جنس به صورت نمودارهای مقایسه ای بررسی کرد. که چند نمونه از ارزیابی عملکرد آموزشی تهیه شده در استان مرکزی جهت اطلاع، به شرح زیر آورده شده است.

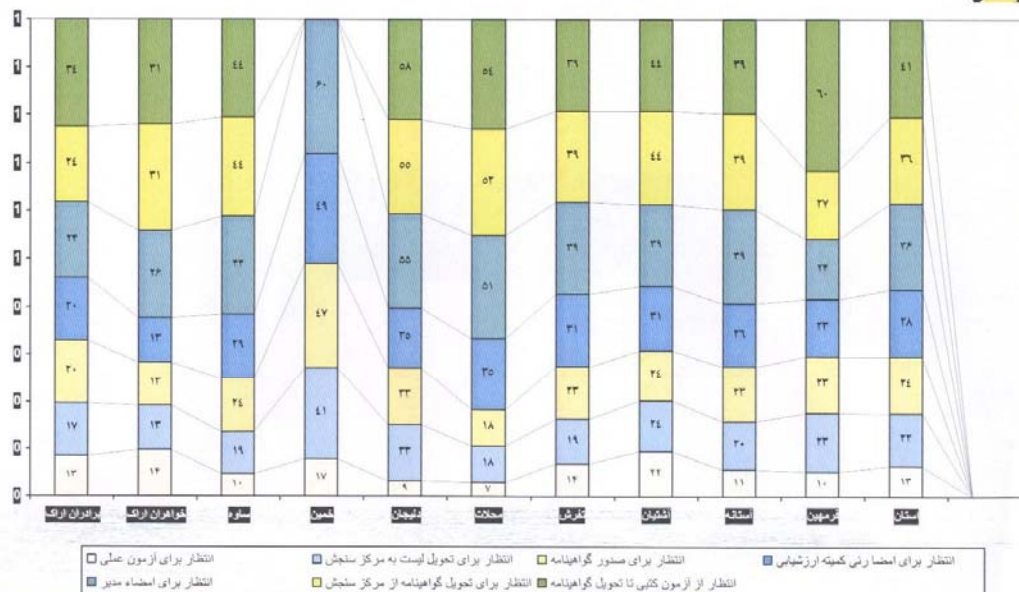
برای بررسی نتایج آزمون عملی نیز پس از وارد کردن نمره عملی در سیستم و با استفاده از نرم افزار excel می توان به همین شیوه ارزیابی عملکرد آموزشی در آزمون کتبی، نمودارهای مقایسه رسم کرد:



نمودار مقایسه ای عملکرد آموزشی بخشهای مختلف در آزمون
اردیبهشت ۱۳۸۸



نمودار نمایش زمان انتظار فرایند های آزمون از آزمون کتبی تا تحویل گواهینامه -ارزشیابی استان-
اردیبهشت ۱۳۸۸



منابع وماخذ:

۱. شریفی، حسن پاشا اصول روان سنجی و روان آزمایی، چاپ دوازدهم بهار ۱۳۸۷، انتشارات رشد
۲. ثرنادایک، رابرت، ال. روان سنجی کاربردی ترجمه هومن، حیدرعلی، دانشگاه تهران، ۱۳۶۹
۳. سیف، علی اکبر، اندازه گیری، سنجش و ارزشیابی آموزشی، ویرایش سوم، چاپ دوازدهم زمستان ۱۳۸۲
۴. مری جی. آلن؟، وندی ام. بن، مقدمه ای بر نظریه اندازه گیری (روانسنجی) ترجمه دلاور، علی، انتشارات سمت، ۱۳۷۴
۵. کرلینگر، فرد، ان. مترجمان: شریفی، حسن پاش و نجفی زند، جعفرقلی، انتشارات آوای نور، جلد اول، ۱۳۷۴
۶. منصور فر کریم، روش های آماری، آذرماه ۱۳۷۱، انتشارات و چاپ دانشگاه تهران.
۷. نشریه آماری ارزیابی عملکرد آموزشی (مراکز آموزش فنی و حرفه ای) استان مرکزی در آزمونهای هماهنگ اردیبهشت ماه ۱۳۸۸ .

ابوالقاسم رافع - آذرماه ۱۳۸۸