

ضرب آلفای کرنباخ و ویژگیهای آن با تأکید بر کاربرد در پژوهشهای اجتماعی

وحید قاسمی*

چکیده

پایایی^۱ ابزارهای سنجش در پژوهشهای اجتماعی یکی از مهمترین موضوعات روش شناختی است چرا که بدون آن نمیتوان به نتایج مطالعات تجربی اعتماد کافی داشت. هماهنگی درونی بین آیتها و ضرب آلفای کرنباخ که بر پایه آن قرار دارد از عمومیترین شیوههای بررسی پایایی ابزارهای سنجش است. در این متن ضمن پرداختن به کلیاتی درباره آلفای کرنباخ، برخی ویژگیهای جزئی تر این ضرب نیز مورد توجه قرار گرفته است. توجه به اینکه پایین بودن مقدار ضرب آلفا لزوماً به معنای عدم پایایی ابزار نیست، توجه به این نکته که آلفای کرنباخ برای ابزارهای تک بعدی مناسب است و در صورت چند بعدی بودن ابزار باید برای هر بعد بنحو جداگانه به محاسبه این ضرب دست زد، ارتباط ضرب آلفای کرنباخ با برخی روشهای دیگر به ویژه روش دو نیمه کردن ابزار سنجش و همچنین دامنه تغییرات این ضرب و شیوههای محاسبه آن از جمله موضوعاتی است که در این مقاله به آن پرداخته شده است.^۲

*گروه علوم اجتماعی دانشگاه اصفهان.

1. Reliability

۲- در جمع آوری دادهها، فرض بر این است که پژوهشگر توجه لازم را به موضوعاتی نظیر همگن سازی شرایط و موقعیت پاسخگویان، صلاحیت پاسخگویان و صلاحیت پرسشگران داشته است و آن را در عمل تدارک دیده است.

واژه‌های کلیدی: اعتبار، پایایی، آلفای کرباخ

مقدمه

پایایی یکی از مهمترین ویژگیها برای ابزارهای سنجشی است که جهت اندازه‌گیری متغیرها یا سازه‌های پنهان طراحی شده‌اند. بدون داشتن یک ابزار پایا نمیتوان به نتایج پژوهش اعتماد کافی داشت و در صورت تکرار پژوهش، نتایج می‌تواند دارای تفاوت معنی‌دار با مرحله اول باشد. هر چند پایایی شرط کافی برای اعتماد و اطمینان به نتایج به بار آمده از یک ابزار سنجش نیست اما شرط ضروری و لازم است. شیوه‌های مختلفی برای سنجش پایایی وجود دارد که عبارتند از: آزمون-آزمون مجدد، استفاده از فرمهای موازی، دو نیمه کردن ابزار سنجش و هماهنگی درونی بین آیتم‌های ابزار سنجش.^۱ برای سنجش هماهنگی درونی بین آیتم‌های مختلف یک ابزار نیز از تکنیک‌های مختلفی میتوان استفاده کرد که استفاده از ضریب آلفای کرباخ یکی از آنهاست. ضریب آلفای کرباخ در علوم مختلف بویژه علوم انسانی و پزشکی برای سنجش پایایی ابزارهای مختلف سنجش نگرش و یا آزمونهای آموزشی و سنجش دانش دارای کاربرد فراوان است. در هر حال با توجه به اینکه سنجش نگرشها بخش مهمی از مطالعات اجتماعی و بویژه پیمایشها را به خود اختصاص داده است، این مقاله با توجه به این نوع از کاربرد ضریب آلفای کرباخ نگاشته شده است. بررسی برخی از مهمترین ابزارهای سنجش نگرش نشان می‌دهد که این ضریب به منظور ساخت ابزار قابل اعتماد مورد استفاده بسیاری از پژوهشگران در حوزه مطالعات اجتماعی بوده است.^۲ پس از طرح ضریب آلفای کرباخ به طور متوسط سالانه ۱۳۱ بار در مقاله‌های علمی به نوشته لی جی. کرباخ در این باره استناد شده است (Cronbach, 2004: 2).

لی جی. کرباخ (Cronbach, 1951) که مهمترین زمینه پژوهشی وی نظریه اندازه‌گیری یا سنجش^۳ می‌باشد، در سال ۱۹۵۱ در مقاله‌ای با عنوان «ضریب آلفا و ساختار درونی آزمونها» این ضریب را به جامعه علمی معرفی کرد. «ضریب آلفا بر پایه هماهنگی درونی آیتم‌ها^۴ با یکدیگر

۱- بر اساس نظر کرلینجر (۱۹۸۶) در حالیکه آزمون - آزمون مجدد بیشتر ویژگی ثبات (Stability) را اندازه می‌گیرد استفاده از فرمهای موازی، دونیمه کردن و هماهنگی درونی گویه‌های به ویژگی دقت (Accuracy) سنجش مرتبط است.

۲- در این رابطه رجوع شود به: میلر، دلبرت. (۱۳۸۰). راهنمای سنجش و تحقیقات اجتماعی (ترجمه هوشنگ نایی). تهران، نشر نی. تاریخ انتشار به زبان اصلی ۱۹۹۱.

طرح شده است» (Cronbach, 1951: 301). ضرب آلفای کرنباخ به عنوان یکی از ضرایب پایایی یا قابلیت اعتماد شناخته می‌شود. این ضریب از عمومی‌ترین ضرایب است که توسط پژوهشگران علوم اجتماعی برای سنجش پایایی ابزارهای مختلف جمع‌آوری داده‌ها مورد استفاده قرار می‌گیرد. مهمترین علت عمومیت استفاده از آن، به احتمال زیاد ناشی از ساده‌تر بودن استفاده از این ضریب در عمل است. با این وجود مباحث طرح شده در کتابهای روش‌شناسی علوم اجتماعی، اغلب به اختصار از کنار آن گذشته‌اند و شاید به همین دلیل - یا دلایل دیگر - برخی ادراکات نادرست از این ضریب در بین دانشجویان علوم اجتماعی وجود دارد. در این مقاله سعی شده است تا با ارایه مثالهایی از متغیرهای جامعه‌شناختی و همچنین تأمل بیشتر در ویژگیهای ضرب آلفای کرنباخ، امکان استفاده صحیح‌تر از آن فراهم شود.

ویژگیهای ضرب آلفای کرنباخ

هر چند ضرب آلفا از طریق فرمول‌های مختلفی قابل محاسبه است اما عمومی‌ترین فرمول مورد استفاده که اولین بار توسط کرنباخ ارائه شد، به شرح زیر است:

$$\alpha = \frac{k}{k-1} \times \left[1 - \frac{\sum S_i^2}{S^2} \right] \quad \text{فرمول (۱)}$$

(Cronbach, 1951: 321; Cronbach, 2004: 9)

که در آن k نماد تعداد آیتم‌ها (گویه، سؤال، معرف و...)، $\sum S_i^2$ نماد مجموع واریانس هر یک از آیتم‌های سازنده ابزار سنجش و S^2 نماد واریانس مجموع نمرات آیتم‌های سازنده ابزار است. S_i^2 جمع ماتریس کوواریانس است.^۱ این ماتریس برای داده‌های صفحه بعد به شرح زیر است:

جدول ۱: ماتریس کوواریانس، جمع و کاربرد آن در محاسبه ضرب آلفای کرنباخ

	گویه اول	گویه دوم	گویه سوم	گویه چهارم
گویه اول	۱/۶۹۷	۱/۴۳۴	۱/۳۷۹	۱/۲۲۸
گویه دوم	۱/۴۳۴	۱/۷۵۲	۱/۳۴۵	۱/۲۳۴
گویه سوم	۱/۳۷۹	۱/۳۴۵	۱/۶۶۱	۱/۰۹۸
گویه چهارم	۱/۲۲۸	۱/۲۳۴	۱/۰۹۸	۱/۳۴۴
$\Sigma = ۲۱/۸۹۰$				

۱- فرمول مربوط به محاسبه ضرب آلفای کرنباخ غیراستاندارد است.

۲- اخذ این ماتریس در بسته نرم‌افزاری SPSS/Windows از طریق Scale-Statistics-Inter-Item Covariances امکان‌پذیر است.

ویژگیهای مختلف ضریب آلفای کرنباخ به لحاظ روش شناختی و همچنین به لحاظ تکنیکی شامل موارد زیر است:

- ۱- ضریب آلفای کرنباخ بر پایه هماهنگی درونی گویه‌ها با یکدیگر قرار دارد.
- ۲- ضریب آلفای کرنباخ تابع دو متغیر «تعداد آیتم‌های ابزار» و «متوسط ضریب همبستگی بین آیتم‌ها»^۱ است.
- ۳- یک ضریب آلفا به تنهایی برای آزمونهایی که هر خوشه از گویه‌ها در حول و حوش وجه ویژه‌ای از پدیده مورد سنجش قرار دارند مناسب نیست (Cronbach, 2004: 20). ضریب آلفای کرنباخ مناسب برای سنجش پایایی ابزارهای تک بعدی^۲ است (SPSS Inc, 1996: 6).
- ۴- مقدار ضریب آلفای کرنباخ برابر با میانگین کلیه ضرایب پایایی ممکن از طریق دو نیمه کردن است (Wuensch, 2004: 4).
- ۵- ضریب آلفای کرنباخ به عنوان یکی از ضرایب محافظه کار شناخته می‌شود به نحوی که ضریب پایایی محاسبه شده از سایر شیوه‌ها (نظیر دو نیمه کردن) مقادیری بیشتر از آلفا را نشان می‌دهد.
- ۶- در عمل، مقدار ضریب آلفای کرنباخ از حداقل $-\infty$ تا حداکثر $+1$ تغییر می‌کند (Nichols, 1999: 1-8).
- ۷- فرمول کرنباخ دارای کاربرد گسترده است و میتوان از آن در سنجش پایایی آزمونهای شناختی دوگزینه‌ای تا مقیاسهای سنجش نگرش چندگزینه‌ای بهره‌بردار (Shavelson, 2003: 381).

یک مثال^۳

ابزاری برای سنجش متغیر «نگرش به دوستی دختر و پسر قبل از ازدواج» تهیه شده است. این ابزار شامل چهار آیتم به شرح جدول ۲ است. به منظور سنجش پایایی ابزار بر اساس روش هماهنگی درونی گویه‌ها و محاسبه ضریب آلفا، گویه‌ها در اختیار ۳۰ نمونه تصادفی از جامعه

1. Inter-Item Correlation

2 Unidimensional Instruments

۳- مثال ارابه شده از پژوهش انجام شده با عنوان «تحلیلی بر روابط دوستانه دانشجویان دختر و پسر در دانشگاه صنعتی اصفهان» که در سال ۱۳۸۳ توسط نگارنده اجرا شده، اخذ گردیده است.

آماري مورد مطالعه گذاشته می شود. امتیاز هر فرد انتخاب شده به عنوان نمونه برای هر یک از گویه ها و همچنین جمع امتیاز هر فرد پاسخگو در جدول ۳ گزارش شده است.

جدول ۲: گویه های سنجش نگرش به دوستی دختر و پسر قبل از ازدواج
و نحوه امتیازدهی به پاسخها

گزینه ها و امتیاز مربوط به آن					گویه	نوع گویه
کاملا مخالف	مخالف	بینابین	موافق	کاملا موافق		
-۲	-۱	۰	+۱	+۲	قبل از ازدواج، داشتن رابطه با جنس مخالف برای هر فردی ضروری و لازم است.	مثبت
-۲	-۱	۰	+۱	+۲	در داشتن رابطه منافع و تجارب زیادی هست که به خاطر آسیبهای احتمالی نباید آن را محدود کرد.	مثبت
+۲	+۱	۰	-۱	-۲	رابطه دوستانه قبل از ازدواج، از عوامل اصلی سست شدن پایه های خانواده در حال حاضر است.	منفی
+۲	+۱	۰	-۱	-۲	هیچ ضرورتی برای رابطه دوستی دختر و پسر قبل از ازدواج وجود ندارد.	منفی

جدول ۳: امتیاز پاسخگویان به هر یک از گویه های طرح شده و مجموع امتیاز از چهار گویه

ردیف	گویه ۱	گویه ۲	گویه ۳	گویه ۴	جمع امتیاز	رتبه	گویه ۱	گویه ۲	گویه ۳	گویه ۴	جمع امتیاز
۱	+۱	+۱	+۲	+۱	+۵	۱۶	+۲	+۲	+۲	+۱	+۷
۲	+۲	+۱	+۱	+۱	+۵	۱۷	-۱	-۲	-۲	-۱	-۶
۳	+۲	+۱	+۱	۰	+۴	۱۸	+۱	+۲	۰	+۱	+۴
۴	+۱	۰	+۱	۰	+۲	۱۹	-۲	-۲	-۲	-۲	-۸
۵	-۱	-۲	-۱	۰	-۴	۲۰	+۱	+۲	+۱	+۱	+۵
۶	-۱	-۱	-۱	-۱	-۴	۲۱	+۱	+۱	+۲	۰	+۴
۷	-۲	-۱	-۱	-۲	-۶	۲۲	-۱	-۱	۰	-۲	-۴
۸	۰	-۱	-۱	-۱	-۳	۲۳	-۱	-۱	-۱	-۱	-۴
۹	-۱	۰	-۲	-۱	-۴	۲۴	+۱	+۱	۰	+۱	+۳
۱۰	+۱	+۲	+۱	+۱	+۵	۲۵	۰	۰	+۱	+۱	+۲
۱۱	+۱	+۱	+۱	+۱	+۴	۲۶	+۲	+۱	+۱	+۱	+۵
۱۲	+۲	+۱	+۲	+۱	+۶	۲۷	۰	-۱	-۱	-۲	-۴
۱۳	-۱	-۲	-۱	-۱	-۵	۲۸	-۱	۰	-۱	۰	-۲
۱۴	+۱	۰	۰	+۱	+۲	۲۹	+۱	+۱	+۱	۰	+۳
۱۵	+۲	+۱	۰	+۱	+۴	۳۰	+۲	+۲	+۲	+۲	+۸

واریانس محاسبه شده برای هر یک از گویه ها و مجموع امتیاز چهار گویه به شرح جدول

شماره ۴ است.

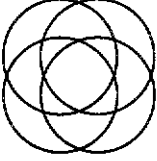
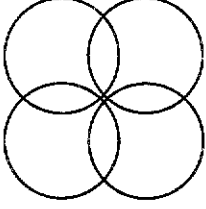
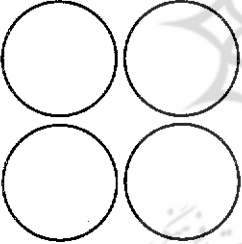
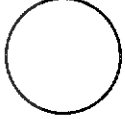
جدول ۴: واریانس محاسبه شده برای هر یک از گویه ها و مجموع نمرات

مقدار	واریانس
۱/۶۹۷	گویه اول s_1^2
۱/۷۵۲	گویه دوم s_2^2
۱/۶۶۱	گویه سوم s_3^2
۱/۳۴۴	گویه چهارم s_4^2
۶/۴۵۴	جمع گویه اول تا چهارم $\sum s_i^2$
۲۱/۸۹۰	مجموع نمره s_i^2

بر اساس واریانس‌های محاسبه شده ضرب آلفای کرنیخ برابر با ۰/۹۴ محاسبه شد. در اغلب منابع در رابطه با مقدار مطلوب برای ضرب آلفا، به نوشته نونالی استناد شده است (Nunally, 1978). بر اساس استدلال وی، برای پایا قلمداد کردن یک ابزار، حداقل مقدار ۰/۷ برای ضرب آلفا لازم است، لذا میتوان ابزار طرح شده را بر اساس شیوه هماهنگی درونی گویه ها، پایا قلمداد کرد.^۱ در هر حال باید توجه داشت که صرف ملاک کمی را نمیتوان دلیل کافی برای پایایی ابزار تلقی کرد و تحلیل کیفی گویه‌ها بر اساس گفتگوهای تخصصی در رابطه با سازه مورد نظر از اهمیت بالایی برخوردار است. در اینجا فرض بر این است که چنین توجه و دقتی توسط پژوهشگر یا گروه پژوهشی صورت گرفته است.

محاسبه ضرب آلفای ۰/۹۴ حاکی از وجود واریانس مشترک (کوواریانس) بالا و در جهت مثبت بین واریانس‌های چهار گویه طرح شده است. وجود واریانس مشترک بالا باعث می شود واریانس مجموع نمرات بسیار بیشتر از جمع واریانس تک تک گویه‌ها شود و در نهایت ضرب به مقدار یک نزدیک شود. میتوان چهار حالت اصلی زیر را در باره واریانس گویه های مختلف در اشکال زیر ترسیم نمود:

۱- در صورت خواست محقق برای کند و کاو بیشتر در داده‌های گردآوری شده میتوان به آزمون این فرضیه صفر که «متغیرهای مورد مطالعه از جامعه آماری یکسانی اخذ شده اند» یا اینکه «میانگین گویه‌های مورد مطالعه با یکدیگر برابرند» دست زد. در این رابطه میتوان از آزمونهای Friedman Chi-Square، Cochran Chi-Square یا F با توجه سطح سنجش متغیرهای مورد مطالعه و نوع توزیع آنها بهره برد. کلیه آزمونهای ذکر شده در بسته نرم افزاری SPSS/Windows در دسترس است.

$s_i^2 \gg \sum_{i=1}^k s_i^2$  <p>کوواریانس بالا - مثبت - و ضریب آلفای نزدیک به یک (ابزار پایا)</p>	$s_i^2 \gg \sum_{i=1}^k s_i^2$  <p>کوواریانس اندک - مثبت^۱ - و ضریب آلفای نزدیک به صفر (ابزار غیر پایا)</p>
$s_i^2 = \sum_{i=1}^k s_i^2$  <p>فقدان کوواریانس و ضریب آلفای برابر صفر (ابزار غیر پایا)</p>	$s_i^2 \ll \sum_{i=1}^k s_i^2$  <p>کوواریانس کامل - مثبت - و ضریب آلفای استاندارد یک (ابزار پایا)</p>

شکل ۱: برخی حالات اصلی نسبت واریانس مجموع نمرات به جمع واریانس تک تک آیتم‌ها

آلفای کرنباخ تابع همبستگی بین آیتم‌ها و تعداد آنها

همانطور که گفته شد ضریب آلفای کرنباخ، شاخصی از هماهنگی درونی در بین آیتم‌ها قلمداد می‌شود. بر این اساس انتظار داریم که هر چه همبستگی بین آیتم‌ها بیشتر شود، مقدار

۱- در قسمت پایانی مقاله درباره حالتی که کوواریانس منفی بین واریانس آیتم‌های مختلف برقرار است، ارایه شده است.

ضرب آلفا افزایش یابد. در این رابطه میتوان این موضوع را ذکر کرد که ضرب آلفای کرنباخ تابع «متوسط ضرایب همبستگی بین آیتم‌ها» است.

همچنین این ضرب تابع تعداد آیتم‌هاست. افزایش تعداد آیتم‌ها، اغلب^۱ بر مقدار ضرب خواهد افزود. تعداد کم آیتم‌ها (مثلاً دو آیتم) باعث می‌گردد که تغییرات کم در نوع پاسخ‌ها به شدت بر ضرب همبستگی بین آیتم‌ها اثر گذارد، لذا ثبات آن کمتر است. در برابر با تعداد زیادتر آیتم برای سنجش یک متغیر (مثلاً ده آیتم)، تغییر اندک در نوع پاسخ‌ها اثر قابل توجهی بر مقدار ضرایب همبستگی نخواهد گذاشت، لذا ابزار پایا تر قلمداد خواهد شد^۲.

بر اساس آنچه ذکر شد ضرب آلفای کرنباخ تابع دو عامل «متوسط ضرایب همبستگی بین آیتم‌ها» و «تعداد آیتم‌ها» است. در این رابطه میتوان فرمول زیر را ارایه کرد:

$$\alpha = \frac{k\bar{r}}{1 + (k-1)r} \quad \text{فرمول (۲)}$$

(SPSS Inc, 1996: 1)

که در آن k نماد تعداد آیتم‌ها و \bar{r} نماد متوسط ضرایب همبستگی بین آیتم‌هاست. در رابطه با مثال ارایه شده ماتریس همبستگی و متوسط آن به شرح زیر است:

جدول ۵: ماتریس همبستگی بین چهار گویه سنجش «نگرش به رابطه دوستی

دختر و پسر قبل از ازدواج»

	گویه اول	گویه دوم	گویه سوم	گویه چهارم
گویه اول	۱			
گویه دوم	۰/۸۳۲۱	۱		
گویه سوم	۰/۸۲۱۷	۰/۷۸۸۴	۱	
گویه چهارم	۰/۸۱۳۱	۰/۸۰۴۶	۰/۷۳۴۸	۱
متوسط ضرایب همبستگی بین گویه‌ها = ۰/۷۹۹۱				

بر اساس فرمول ارایه شده مقدار ضرب آلفا برابر با ۰/۹۴ محاسبه می‌شود.

۱- قید اغلب مربوط به شرط طرح مناسب گویه‌هایی است که به لحاظ نظری با سایر گویه‌ها مرتبط باشند.
 ۲- این وضعیت در رابطه با تعداد افراد مورد مطالعه برای سنجش پایایی ابزار نیز صادق است. هر چند در فرمول ارایه شده ضرب آلفا تابع تعداد نمونه مورد مطالعه نیست اما کمی تعداد نمونه مورد مطالعه (بمنزله مثال ۱۰ نفر) باعث می‌شود که پاسخگویی با دقت کم تنها یک یا دو پاسخگو بر مقدار واقعی ضرب اثر نامطلوب گذارد. افزایش تعداد نمونه با صلاحیت برای پاسخگویی به سؤالات منجر به برآورد دقیق‌تر و واقعی‌تری از ضرب آلفا می‌شود.

تک بعدی بودن ابزار سنجش

اگر متغیر مورد سنجش یک سازه تک بعدی است، انتظار می‌رود که همبستگی بالایی بین آیت‌ها وجود داشته باشد، لذا مقدار ضریب آلفا بالا خواهد بود. اما چنانچه متغیر مورد مطالعه یک سازه چند بعدی است نمیتوان انتظار داشت که بین آیت‌های متعلق به ابعاد مختلف همبستگی بالایی وجود داشته باشد و در این حالت مقدار ضریب آلفا، از مقدار $+1$ فاصله خواهد گرفت و احتمالا به پایین‌تر از مقدار 0.7 کاهش خواهد یافت. فرض کنید که مایل هستیم تغییری به نام «دین داری» را اندازه‌گیری کنیم. ابتدا چهار آیت به شرح زیر در نظر گرفته ایم:

جدول ۶: گویه‌های طرح شده برای سنجش دین داری و نحوه امتیاز دهی به پاسخ‌ها

گزینه‌ها و امتیاز مربوط به آن					گویه	نوع گویه
کاملا مخالف	مخالف	بینابین	موافق	کاملا موافق		
۱	۲	۳	۴	۵	به وجود خداوند یکتا باور دارم.	مثبت
۵	۴	۳	۲	۱	باور ندارم که فرشتگان وجود داشته باشند.	منفی
۱	۲	۳	۴	۵	نماز را مرتب می‌خوانم.	مثبت
۵	۴	۳	۲	۱	در گرفتن روزه چندان مصمم نیستم.	منفی

۶۰ نفر پاسخگو که به طور تصادفی از جامعه آماری تعریف شده انتخاب شده‌اند، به چهار گویه فوق پاسخ داده‌اند. ماتریس همبستگی، واریانس هر یک از گویه‌ها و واریانس مجموع نمرات به شرح جداول ۶ و ۷ است:

جدول ۷: واریانس محاسبه شده برای هر یک از گویه‌ها و مجموع نمرات

مقدار	واریانس
۰/۸۲۳	گویه اول S_1^2
۰/۵۷۵	گویه دوم S_2^2
۰/۸۰۰	گویه سوم S_3^2
۰/۹۹۳	گویه چهارم S_4^2
۳/۱۹۱	جمع گویه اول تا چهارم $\sum S_i^2$
۶/۳۲۶	مجموع نمره S_i^2

جدول ۸: امتیاز پاسخگویان به هر یک از گویه های طرح شده و مجموع امتیاز از چهار گویه

ردیف	گویه ۱	گویه ۲	گویه ۳	گویه ۴	جمع امتیاز	رتبه	گویه ۱	گویه ۲	گویه ۳	گویه ۴	جمع امتیاز
۱	۵	۵	۲	۱	۱۳	۱۶	۴	۵	۲	۲	۱۳
۲	۳	۴	۲	۱	۱۰	۱۷	۳	۴	۳	۳	۱۳
۳	۳	۴	۲	۱	۱۰	۱۸	۴	۵	۳	۴	۱۶
۴	۴	۵	۳	۲	۱۴	۱۹	۴	۵	۳	۴	۱۶
۵	۳	۴	۳	۲	۱۲	۲۰	۴	۳	۴	۳	۱۴
۶	۳	۴	۳	۳	۱۳	۲۱	۵	۳	۲	۱	۱۱
۷	۳	۴	۲	۱	۱۰	۲۲	۵	۵	۵	۴	۱۹
۸	۴	۵	۲	۲	۱۳	۲۳	۵	۵	۴	۳	۱۷
۹	۴	۵	۳	۳	۱۵	۲۴	۴	۴	۲	۱	۱۱
۱۰	۳	۴	۳	۳	۱۳	۲۵	۴	۴	۳	۲	۱۳
۱۱	۴	۵	۳	۲	۱۴	۲۶	۴	۴	۳	۲	۱۳
۱۲	۳	۵	۴	۳	۱۵	۲۷	۵	۵	۲	۱	۱۳
۱۳	۲	۳	۲	۱	۸	۲۸	۴	۵	۱	۳	۱۳
۱۴	۲	۳	۲	۲	۹	۲۹	۵	۵	۱	۳	۱۴
۱۵	۲	۳	۲	۱	۸	۳۰	۴	۵	۲	۲	۱۳

ضرب آلفای کرنباخ با در نظر گرفتن چهار گویه برای سنجش متغیر دین داری برابر با ۰/۶۶۱ محاسبه شده است. با توجه به محتوای هر یک از گویه ها، چنین به نظر می رسد که این چهار گویه، شاید در دو گروه دوتایی قابل طبقه بندی هستند. به نحوی که میتوان گفت دو گویه اول متغیر «باورهای دینی» و دو گویه دوم «انجام مناسک دینی» را اندازه گیری می کنند. با استفاده از تحلیل عاملی میتوان این موضوع را به آزمون گذاشت. نتیجه تحلیل عاملی نشان می دهد که میتوان دو گویه اول را به عنوان یک عامل یا بعد و دو گویه دیگر را نیز به عنوان عامل یا بعد دوم در نظر گرفت.

جدول ۹: نتیجه تحلیل عاملی برای تأیید ابعاد دو گانه نهفته در چهار گویه طرح شده

Factor Analysis

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.011	50.266	50.266	2.011	50.266	50.266	1.633	40.822	40.822
2	1.177	29.427	79.694	1.177	29.427	79.694	1.555	38.872	79.694
3	.517	12.929	92.622						
4	.295	7.378	100.000						

Extraction Method: Principal Component Analysis.

Rotated Component Matrix^a

	Component	
	1	2
V1	.877	4.62E-02
V2	.881	.176
V3	-2.90E-02	.905
V4	.293	.838

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

بر اساس نتایج به دست آمده، این راه حل در برابر محقق قرار می گیرد که ضریب آلفای کربناخ را برای هر یک از ابعاد دو گانه فوق بنحو جدا گانه محاسبه نماید (بحث تک بعدی بودن ضریب آلفای کربناخ). واریانس مجموع نمرات گویه های اول و دوم و همچنین واریانس مجموع نمرات گویه سوم و چهارم در جدول زیر گزارش شده است:

جدول ۱۰: واریانس محاسبه شده برای هر یک از گویه ها و مجموع نمرات

مقدار	واریانس
۲/۲۰۲	مجموع گویه اول و دوم
۲/۷۸۶	مجموع گویه سوم و چهارم

بر این اساس محاسبه ضریب آلفا برای هر یک از ابعاد دو گانه تعریف شده برابر ۰/۷۳ و ۰/۷۱ محاسبه می شود که نتیجه قابل قبول تری نسبت به وضعیت قبلی است. در واقع، در حالت قبل (تک بعدی در نظر گرفتن یک ابزار دو بعدی) محاسبه ضریب پایایی، با زیربرآورد همراه بود.

رابطه ضرب آلفا با روش دو نیمه کردن

یکی از روش‌های بررسی پایایی ابزار سنجش، روش دو نیمه کردن است. دو نیمه کردن آیت‌ها به طرق مختلفی صورت می‌گیرد. گزینش تصادفی، گزینش زوج و فرد و همچنین گزینش قضاوتی به منظور برابر در نظر گرفتن دو نیمه از جمله این موارد است. در هر حال با توجه به تعداد گویه‌ها، تعداد ترکیب‌های ممکن برای گویه‌های دو نیمه مشخص می‌شود. ضرب آلفای کرنباخ برابر با متوسط ضرایب پایایی کلیه حالات دو نیمه کردن است. برای دو گویه، تنها ۱ ترکیب، برای چهار گویه، ۳ ترکیب و برای ده گویه ۱۲۶ ترکیب ممکن وجود دارد. تعداد ترکیب‌های ممکن برای دو نیمه تعریف شده از رابطه زیر به دست می‌آید:

$$C = \frac{0.5 \times k!}{(k/2)!} \quad \text{فرمول (۳)}$$

به عنوان مثال برای چهار گویه که ۳ ترکیب امکان پذیر است، ترکیب‌های ممکن عبارتند از:

گویه‌های نیمه اول ^۱	گویه‌های نیمه دوم
۱-۲	۳-۴
۱-۳	۲-۴
۱-۴	۲-۳

ضرب گاتمن و ضرب اسپیرمن- براوان از جمله ضرایبی هستند که بر پایه تکنیک دو نیمه کردن ابزار سنجش قرار دارند و بر اساس رابطه‌های زیر محاسبه می‌شوند:

$$G = 2 \times \left(1 - \frac{s_A^2 + s_B^2}{s_i^2} \right) \quad \text{فرمول (۴)}$$

Source: (Wuensch, 2004: 4)

که در آن G نماد ضرب گاتمن، σ_A^2 نماد واریانس مجموع نمرات آیت‌های نیمه اول، σ_B^2 نماد واریانس مجموع نمرات آیت‌های نیمه دوم و σ_i^2 واریانس مجموع نمرات کلیه آیت‌ها

$$r_{sb} = \frac{2 \times r_s}{1 + r_s} \quad \text{فرمول (۵)} \quad \text{می‌باشد.}$$

که در آن r_{sb} نماد ضرب اسپیرمن- براوان و r_s برابر با ضرب اسپیرمن است.

۱- با توجه به مقارن بودن ضرایب همبستگی اسپیرمن یا گاتمن، اینکه کدام نیمه را اول و کدامیک را دوم در نظر بگیریم، تأثیری در نتایج نخواهد گذاشت.

با در نظر گرفتن مثال مربوط به سنجش متغیر «نگرش به رابطه دوستی بین دختر و پسر قبل از ازدواج» واریانس‌های محاسبه شده برای هر یک از حالات ممکن و ضرایب گاتمن مربوط به شرح جدول زیر می باشد:

جدول ۱۱: ضرایب پایایی گاتمن و اسپیرمن-براون بر پایه تکنیک دو نیمه کردن ابزار سنجش

شماره گویه ها	واریانس		ضریب همبستگی		ضریب پایایی		
	نیمه اول	نیمه دوم	اسپیرمن	گاتمن	اسپیرمن-براون	گاتمن	
۱-۲	۲/۵۱۳	۲/۲۸۰	۰/۸۴۰	۰/۹۴۷۷	۰/۹۵۰۱	۰/۹۴۷۷	
۱-۳	۲/۴۷۳	۲/۳۵۹	۰/۷۹۷	۰/۹۳۲۸	۰/۹۳۳۳	۰/۹۳۲۸	
۱-۴	۲/۳۴۴	۲/۴۷۰	۰/۸۲۱	۰/۹۴۰۳	۰/۹۴۱۰	۰/۹۴۰۳	
واریانس مجموع ۴ گویه = ۴/۶۷۹					۰/۹۴	۰/۹۴	

میانگین محاسبه شده برای ضرایب گاتمن و اسپیرمن-براون در کلیه حالات دو نیمه کردن ممکن برابر با ضریب آلفای کربناخ است، لذا از آنجا که در هنگام دو نیمه کردن این انتظار وجود دارد که محقق بهترین ترکیب از آیت‌ها را برای هر نیمه در نظر گیرد، اغلب ضریب آلفا به عنوان ضریبی محافظه کار تر در برآورد پایایی نسبت به ضرایب گاتمن و اسپیرمن-براون شناخته می شود.

دامنه تغییرات ضریب آلفای کربناخ

در نظریه کلاسیک، کل واریانس نمرات مشاهده در یک تست (σ_X^2) بعنوان مجموع واریانس نمرات واقعی (σ_T^2) و واریانس خطا (σ_E^2) تعریف شده است (Lord & Novick, 1968).

همچنین پایایی یک ابزار تعریف شده است به عنوان:

۱. توان دوم ضریب همبستگی بین نمرات مشاهده شده و نمرات واقعی ρ_{XT}^2
۲. نسبت واریانس نمرات واقعی به واریانس نمرات مشاهده شده σ_T^2 / σ_X^2
۳. یک منهای نسبت واریانس خطا به واریانس نمرات مشاهده شده $1 - (\sigma_E^2 / \sigma_X^2)$

به بیان دیگر:

$$\rho_{XT}^2 = \sigma_T^2 / \sigma_X^2 = 1 - (\sigma_E^2 / \sigma_X^2)$$

یک ویژگی اساسی در تعریف ضریب پایایی این است که نسبت واریانس - مطابق با تئوری - بین صفر تا یک است. اما در عمل ما با نمرات خطا و واقعی مشاهده نشده‌ای روبرو هستیم که وقتی از تئوری وارد عمل می‌شویم، کوشش ما برای تخمین پایایی می‌تواند با نتایج غیرقابل انتظاری روبرو شود. در عمل مقادیر ممکن برآورد پایایی از حداقل $-\infty$ تا حداکثر $+1$ تغییر می‌کند. با اینکه جمله اول از فرمول ضریب آلفای کرنباخ (فرمول ۱) همیشه مثبت است، اما کل نتیجه می‌تواند منفی باشد تنها در صورتی که:

$$\sum \sigma_i^2 / \sigma_X^2 > 1$$

یا به عبارت دیگر:

$$\sum \sigma_i^2 > \sigma_X^2$$

یعنی آلفا در صورتی می‌تواند منفی باشد که مجموع واریانس تک تک گویه‌ها، بزرگتر از واریانس کل مقیاس باشد. از آنجا که واریانس مجموعه‌ای از متغیرهای تصادفی برابر است با مجموع واریانس‌های هر یک از متغیرها با اضافه دو برابر مجموع کوواریانس آنها (Hays, 1981) و از آنجا که نمره هر مقیاس، مجموع نمره هر یک از گویه‌های منفرد است، واریانس مقیاس می‌تواند به صورت زیر بیان شود:

$$\sigma_X^2 = \sum \sigma_i^2 + \sum \sum \sigma_{ij}$$

که در آن σ_{ij} نماد کوواریانس بین گویه i ام و j ام است و دو علامت سیگما برای تمام ترکیب‌هایی که در آن $i \neq j$ است. بنابراین می‌توانیم شرط لازم و کافی برای منفی شدن ضریب آلفا را به صورت زیر بیان کنیم:

$$\sum \sigma_i^2 > \sum \sum \sigma_{ij}$$

به طور مختصر، ضریب آلفا منفی خواهد بود هر گاه دو برابر مجموع کوواریانس آیت‌ها منفی باشد و به طور ساده‌تر ضریب آلفا منفی خواهد بود هر گاه میانگین کوواریانس درمیان آیت‌ها منفی باشد. در این باره به مثال زیر دقت کنید:

جدول ۱۲: امتیاز ۲۵ نفر پاسخگو به چهار گویه فرضی که میانگین کوواریانس آنها منفی است

ردیف	گویه ۱	گویه ۲	گویه ۳	گویه ۴	جمع امتیاز	ردیف	گویه ۱	گویه ۲	گویه ۳	گویه ۴	جمع امتیاز
۱	۱	۲	۴	۳	۱۰	۱۴	۵	۴	۱	۲	۱۲
۲	۳	۳	۴	۳	۱۳	۱۵	۵	۴	۲	۱	۱۲
۳	۳	۴	۳	۳	۱۳	۱۶	۵	۴	۲	۱	۱۲
۴	۳	۴	۳	۲	۱۲	۱۷	۴	۳	۳	۳	۱۳
۵	۲	۳	۳	۲	۱۰	۱۸	۳	۳	۳	۳	۱۲
۶	۲	۲	۳	۲	۹	۱۹	۳	۲	۴	۳	۱۲
۷	۴	۳	۳	۲	۱۲	۲۰	۲	۱	۴	۴	۱۱
۸	۴	۳	۲	۱	۱۰	۲۱	۱	۲	۳	۳	۹
۹	۴	۵	۲	۲	۱۳	۲۲	۱	۲	۳	۳	۹
۱۰	۵	۵	۲	۱	۱۳	۲۳	۱	۲	۳	۲	۸
۱۱	۴	۵	۲	۱	۱۲	۲۴	۲	۲	۲	۲	۸
۱۲	۴	۵	۱	۱	۱۱	۲۵	۲	۳	۲	۲	۹
۱۳	۳	۴	۱	۲	۱۰						

ماتریس کوواریانس و همبستگی بین چهار گویه فرضی فوق به شرح جداول ۱۰ و ۱۱

است.

جدول ۱۳: ماتریس کوواریانس بین چهار گویه فرضی

	گویه اول	گویه دوم	گویه سوم	گویه چهارم
گویه اول	۱/۷۹۰			
گویه دوم	۱/۱۵۸	۱/۳۳۳		
گویه سوم	-/۶۵۰	-/۷۰۸	۰/۸۳۳	
گویه چهارم	-/۶۳۳	-/۵۸۳	۰/۵۴۲	۰/۶۶۷
متوسط کوواریانس بین گویه‌ها = -/۱۴۶				

جدول ۱۴: ماتریس همبستگی بین چهار گویه فرضی

	گویه اول	گویه دوم	گویه سوم	گویه چهارم
گویه اول	۱			
گویه دوم	۰/۷۵۰	۱		
گویه سوم	-/۵۳۲	-/۶۷۲	۱	
گویه چهارم	-/۵۸۰	-/۶۱۹	۰/۷۲۷	۱
متوسط همبستگی بین گویه‌ها = -/۱۵۴				

محاسبه ضریب آلفا با استفاده از بسته نرم افزاری SPSS/Windows مقدار ۰/۸۱۲- را نشان می‌دهد. برخورد با یک ضریب منفی برای آلفای کرنباخ می‌تواند ناشی از انجام برخی خطاها به لحاظ تکنیکی یا روشی باشد. در صورت برخورد کردن با یک ضریب منفی برای آلفای کرنباخ باید به موارد زیر توجه شود:

۱- اولین چیزی که باید چک شود این است که آیا در کدگذاری داده‌ها اشتباهی رخ داده است یا خیر؟ یک اشکال عمومی این است که مقیاس شامل برخی گویه‌هایی است که در جهت متضادی جمله بندی شده‌اند و محقق فراموش کرده است که آنها را به طور مناسب باز کدگذاری^۱ کند و در نتیجه کوواریانس منفی محاسبه شده است، در حالی که واقعا مثبت است.

۲- احتمال دیگر که در موارد با نمونه‌های کم و همچنین تعداد کم گویه‌ها رخ می‌دهد، این است که در حالی که کوواریانس در بین گویه‌ها مثبت است، خطای نمونه‌گیری باعث ایجاد یک کوواریانس منفی در یک نمونه خاص شده است.

۳- در نهایت ممکن است محقق در تهیه گویه‌هایی که پدیده یکسانی را اندازه‌گیری می‌کنند دچار اشتباه شده باشد و گویه‌های تعریف شده واقعا دارای کوواریانس منفی باشند و بنابراین مجموعه گویه‌ها تشکیل دهنده مقیاس منفردی که پدیده یکسانی را بسنجند، نیستند.

بدترین وضعیت ممکن برای تعریف یک ابزار که پدیده واحدی را بسنجد، این است که ضریب آلفای کرنباخ مقدار ۰۰- را به خود بگیرد. در یک مورد ساده‌تر، این وضعیت هنگامی که اتفاقی می‌افتد که ابزار شامل دو آیتم است که ضریب همبستگی بین آنها یک ضریب کامل در

جهت منفی است. واریانس دو گویه با یکدیگر برابر است، لذا واریانس مجموع نمرات نیز برابر با صفر می شود.

ضریب آلفای استاندارد و غیراستاندارد

محاسبه ضریب آلفای غیر استاندارد بر پایه واریانس و کوواریانس قرار دارد (فرمول ۱) و محاسبه ضریب آلفای استاندارد بر پایه ضریب همبستگی (فرمول ۲) یا نمرات استاندارد قرار دارد. در ابتدا ذکر این نکته اهمیت دارد که «نمی‌توان به طور مطلق یکی از انواع ضرایب را نسبت به دیگری برتر یا مطلوبتر دانست» (Yu, 1993: 3). هر چند این دو مفهوم با یکدیگر رابطه دارند اما دقیقاً یکسان نیستند. به لحاظ کاربردی میتوان چنین قضاوت کرد که در برخی شرایط استفاده از هر یک از دو روش فوق نتایج یکسان یا تقریباً یکسانی به بار خواهد آورد و در برخی شرایط نتایج تا حد قابل توجهی متفاوت خواهند بود.

در اغلب ابزارهای سنجش در پژوهشهای اجتماعی که برای اندازه گیری یک متغیر خاص طراحی شده‌اند، گزینه‌های مختلف هر آیتم «اغلب» دامنه تغییرات یکسانی داشته و واریانس گویه‌های مختلف به یکدیگر نزدیک است. به عنوان مثال استفاده از طیف لیکرت که امتیاز هر گویه بین ۱ تا ۵ تغییر می کند یا سؤالات شناختی که امتیاز هر پاسخ (پاسخ غلط) یا ۱ (پاسخ صحیح) است. سایر طیفهای مورد استفاده برای سنجش متغیرهای پنهان نظیر طیف بوگاردوس یا افتراق معنایی، نیز اغلب ۵ گزینه‌ای هستند. در این حالت ضرایب آلفای استاندارد و غیراستاندارد به یکدیگر نزدیک خواهد بود.

در حالت دوم - که کمتر اتفاق می افتد - دامنه تغییرات آیتم‌های مختلف متفاوت است. مثلاً برخی از آیتم‌ها ۰ تا ۱ و برخی از آیتم‌ها ۱ تا ۵. در چنین حالتی که واریانس آیتم‌های مختلف تفاوت‌های زیادی در مقایسه با یکدیگر دارند، نتایج روش استاندارد و غیراستاندارد متفاوت خواهد شد. «در این حالت استفاده از ضریب استاندارد ترجیح داده می‌شود» (Renaldo & Santos, 1996: 6).

در هر حال از آنجا که استفاده از روش استاندارد در حالت‌های مختلف (آیتم‌های دارای گزینه‌های با دامنه تغییرات یکسان یا متفاوت) نتایج یکسانی به بار می‌آورد و همچنین اینکه امکان مقایسه پایایی ابزارهای مختلفی که برای سنجش متغیرهای یکسانی توسط گروه‌های پژوهش طرح شده‌اند را امکان پذیر می‌سازد، ترجیح داده می‌شود.

نتیجه گیری

در استفاده از ضرب آلفای کرنباخ باید به برخی نکات مهم توجه شود:

- پایین بودن ضرب آلفای کرنباخ را نباید دلیل کافی برای ناپایا بودن ابزار سنجش تلقی کرد.
- بالا بودن ضرب آلفای کرنباخ به معنای مناسب بودن ابزار طراحی شده به لحاظ نظری نبوده و چه بسا با برخی ملاحظات بتوان بدون تغییر در ابزار به ضرایب پایایی بالاتری دست یافت.
- از آنجا که نه بالا بودن ضرب آلفای کرنباخ (بیشتر از ۰/۷) ضرورتاً به معنای پایایی مطلوب و نه پایین بودن آن ضرورتاً به معنای ضعف پایایی واقعی ابزار سنجش است، میتوان گفت استفاده از تحلیل نظری و کیفی در کنار استفاده از تکنیکهای کمی نظیر ضرب آلفای کرنباخ ضرورت دارد.^۱
- توجه به ماهیت ضرب آلفای کرنباخ به عنوان تکنیکی برای سنجش هماهنگی درونی آیتم‌های یک ابزار ضروری است.
- چنانچه متغیر مورد سنجش توسط ابزار یک متغیر چند بعدی باشد با توجه به اینکه این انتظار می‌رود که همواره آیتم‌های سازنده یک بعد با همدیگر نسبت به آیتم‌های سایر ابعاد، هماهنگ‌تر باشند، لذا محاسبه این ضرب برای هر یک از ابعاد بنحو جداگانه برای دستیابی به ضرب پایایی واقعی ضروری است.
- منفی بودن ضرب آلفای کرنباخ در صورت فقدان مسائیل تکنیکی در محاسبه ضرب به معنای نقصان جدی در طرح ابزار سنجش تفسیر خواهد شد.
- هر چند ضرایب آلفای استاندارد و غیراستاندارد در مقایسه با یکدیگر دارای برتری مطلق نیستند اما به برخی دلایلی که در متن ذکر شد، استفاده از ضرب استاندارد توصیه می‌شود.
- سنجش کمی پایایی با استفاده از ضرب آلفای کرنباخ به لحاظ روش شناختی روبنای تحلیل کیفی گویه‌ها و پس از آن قرار دارد. در این رابطه بحث و بررسی کارشناسی گویه‌ها در گروه‌های گفتگوی تخصصی به منظور تدارک گویه‌های مناسب برای دستیابی به یک ابزار سنجش پایا می‌تواند بسیار راهگشا باشد.

۱- در هر حال توجه خواننده گرامی را به این نکته توجه می‌دهد که تمرکز مقاله حاضر نه پیرامون بحث پایایی یا قابلیت اعتماد (Reliability) بلکه درباره ضرب آلفای کرنباخ به عنوان یک تکنیک کمی در این رابطه بوده است.

منابع

- ۱- کرلینجر، فردان، مبانی پژوهش، در علوم رفتاری (ترجمه حسن پاشا شریفی و جعفر نجفی زند)، تهران، آوای نور، تاریخ انتشار به زبان اصلی ۱۹۸۶، ۱۳۸۲.
- ۲- میلر، دلبرت، راهنمای سنجش و تحقیقات اجتماعی (ترجمه هوشنگ نایی)، تهران، نشر نی، تاریخ انتشار به زبان اصلی ۱۹۹۱، ۱۳۸۰.
3. Cortina, J. M. (1993). What is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology*, No:78, PP. 98-104.
4. Cronbach, Lee J. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, No: 16, PP. 297-334.
5. _____ (2004). My Current Thoughts on Coefficient Alpha and Successor Procedures. *Educational and Psychological Measurement*, 64, 391-418.
6. Hays. W. L. (1981). *Statistics*. (3rd ed.) Holt, Rinehart and Winston.
7. Nichols, David P. (1999). My Coefficient Alpha is Negative. SPSS Inc, *Keywords*, Number 68.
8. Nunally, J. C. (1978). *Psychometric Theory* (2nd ed.). New York: McGraw-Hill.
9. Reynaldo, J.; and Santos, A. (1999). *Cronbach's Alpha a Tool for Assessing the Reliability of Scales*. Texas: A & M University.
10. Shavelson, Richard J. (2003). *Lee J. Cronbach*. Proceedings of the American Philosophical Society, Vol: 147, No: 4, December 2003.
11. SPSS Inc, Statistical Consulting Services. (1996). *What Does Cronbach's Alpha Mean?* [on-line]. Available at www.ats.ucla.edu/stat/spss/alpha.html.
12. Wuensch, Karl L. (2004). *Cronbach's Alpha and Maximized Lambda4*. Core.ecu.edu\psyc\wuenschk\mu\alpha. Doc.
13. Yu, H. C. (1993). *An Introduction to Computing and Interpreting Cronbach Coefficient Alpha*. USA: Arizona State University.