

پایایی

تعریف پایایی

پایایی یک وسیله اندازه گیری به دقت آن اشاره می کند. یک آزمون در صورتی دارای پایایی است که اگر آن را در یک فاصله زمانی کوتاه چندین بار به گروه واحدی از افراد بدهیم نمرات حاصل از این چندین بار اجرا نزدیک به هم باشند.

رابطه بین پایایی و روایی

رابطه بین روایی و پایایی از این قرار است که یک آزمون باید پایا باشد تا بتواند روا باشد. اگر آزمونی در هر بار اجرا در مورد تعدادی دانش آموز نتایج مختلفی بدست بدهد آن آزمون یک آزمون پایا نخواهد بود. پس برای این که یک آزمون روا باشد باید پایا باشد، یعنی پایایی شرط روایی است، اما روایی بر پایایی ضروری نیست.

انواع سنجش های پایایی

روش مورد استفاده برای سنجش پایایی یک ابزار (آزمون یا پرسش نامه) به دلیل عملکرد و طبیعت ابزار و چگونگی استفاده از آن در زمینه مورد استفاده متفاوت می باشد. گاهی هدف تعیین ضریب پایایی به معنای تعیین توافق درونی است که از روش های آلفای کرونباخ، کادر ریچارسون و دو نیمه کردن می توان استفاده کرد. اما چنانچه بخواهیم ثبات ابزار را مورد ارزیابی قرار دهیم می توان از روش آزمون باز آزمون و یا از روش فرم های موازی استفاده کرد. همچنین گاهی هدف از تعیین ضریب پایایی تعیین توافق بین مشاهده گران یا مصاحبه کنندگان و یا ارزیابان می باشد.

سنجش توافق درونی در پایایی

۱. روش القای کرونباخ (Cronbach's Alpha)

کرونباخ در سال ۱۹۵۱ در دانشگاه استنفورد روش آماری ضریب آلفا را برای حل مشکل تعیین پایایی آزمون های چند سوالی ابداع کرد. این روش معمول ترین ضریب پایایی ثبات داخلی است که در بیشتر مطالعات از آن استفاده می شود و معرف میزان تناسب گروهی از آیتم هایی است که یک سازه را می سنجند.

مقدار آلفا باید حداقل برابر با ۰/۷ و یا بیشتر باشد تا یک سوال بتواند در یک ابزار باقی بماند و بسیاری از محققان نقطه ۰/۸ را برای باقی ماندن یک سوال در ابزار ضروری می دانند. پایایی برابر با ۰/۷ ثبات ۷۰ درصدی امتیازات یا نمرات به دست آمده از یک ابزار را نشان می دهد و روشی برای محاسبه ثبات درونی بر اساس میانگین همبستگی درونی سوالات می باشد و هنگامی که مقیاسی از نوع مقیاس لیکرت در ابزار اطلاعات به کار رود، محاسبه مقدار ضریب آلفا جهت تعیین پایایی ثبات درونی امری ضروری است.

مقدار عددی آلفای کرونباخ برابر با متوسط مقدار ضریب پایایی است که کل ترکیب های ممکن آیتم های ابزار، پس از آن که به دو نیمه تقسیم شود به دست می آید. روش آلفای کرونباخ یکی از روش های سنجش پایایی است که فقط به انجام یک بار آزمون نیاز دارد تا برآوردی از پایایی آزمون را فراهم کند.

در کل ضریب پایایی آلفای کرونباخ هنگامی مفید است که سوالات یک ابزار به صورت صحیح - غلط امتیاز بندی نشده باشند و جهت سنجش ثبات درونی سوالات در مقیاس لیکرت به کار می رود. در سال ۲۰۰۳ مقدار آلفای بیش از ۰/۹ را عالی، ۰/۸ - ۰/۹ را خوب، ۰/۷ - ۰/۸ را قابل قبول، ۰/۷ - ۰/۶ را قابل بحث، ۰/۵ - ۰/۶ را ضعیف و کمتر از ۰/۵ را غیر قابل قبول معرفی کردند.

فرمول تعیین آلفای کرونباخ به قرار زیر است:

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum \delta_i^2}{\delta_i^2} \right)$$

در این فرمول k تعداد آیتم های ابزار δ_i^2 و همچنین i و δ_i^2 واریانس کل ابزار یا تست که از مجموع کل آیتم ها بدست آمده است، می باشد.

۲. پایایی دو نیمه آزمون (Split half reliability)

در این روش یک آزمون به دو نیمه تقسیم شده و رابطه همبستگی بین دو نیمه آزمون سنجیده می شود. این آزمون معمول برای متغیرهای دو حالتی که کد صفر به پاسخ های غلط و کد یک به پاسخ های صحیح داده می شود به کار می رود. در این روش دو نیمه آزمون باید در محتوا و سختی سوالات با هم مشابه باشند و از آن جا که تعداد سوالات ابزار به دو قسمت تقسیم شده است می بایست روش ضریب همبستگی به کار رود که با استفاده از ضریب اسپیرمن و گاتمن محاسبه می شود.

ضریب اسپیرمن براون فرمی از سنجش پایایی دو نیمه آزمون است که برای برآورد پایایی کل آزمون به کار می رود. همبستگی پیرسون بین دو نیمه آزمون، پایایی نیمه هایی از ابزار یا مقیاس را برآورد می کند اما ضریب پیشگویی اسپیرمن، پایایی کل آزمون را بر اساس نیمه های آن برآورد می کند. مقدار این ضریب بیشتر از ضریب پایایی نیمه های آزمون (IX و IX) است و معمولاً مساوی با دو برابر ضریب نیمه آزمون تقسیم بر مقدار یک به اضافه پایایی نیمه دیگر آزمون خواهد بود. ضریب پایایی اسپیرمن براون با فرمول زیر محاسبه می شود:

$$SB_{prxx} = \frac{2rx_1x_2}{1+rx_1x_2}$$

در بین سنجش ها d پایایی دو نیمه آزمون، ضریب پایایی اسپیرمن شدیداً تحت تأثیر ترتیب و نحوه قرار گرفتن سوالات ابزار مورد استفاده در هر یک از دو نیمه آزمون می باشد. این فرآیند در این روش باید به شکل کاملاً دقیقی، تصادفی انجام شود و تخصیص تصادفی سوالات بین دو نیمه آزمون باید تساوی واریانس بین دو فرم را تأمین نماید. اما همیشه این امر باید توسط محقق چک شود؛ زیرا هیچ گاه این امر تضمین نمی شود و مقدار پایایی از یک اندازه گیری به اندازه گیری بعدی متفاوت خواهد بود.

ضریب گاتمن فرض می کند که پایایی نیمه اول و دوم (دو نیمه فرم) یکسان است و به تساوی واریانس های بین دو نیمه آزمون نیازی ندارد. بهترین نیمه هنگامی ایجاد می شود که هر نیمه سوالاتی را که بیشترین همبستگی درونی دارند، شامل شود. به عبارتی زمانی که ابزاری به دو نیمه تقسیم شود به طوری که هر نیمه سوالاتی را که شدیداً وابسته به همان گروه هستند و هیچ گونه وابستگی بین دو گروه ندارند را شامل شود، ضریب دو نیمه گاتمن کم ترین مقدار ممکن را خواهد گرفت و زمانی که سوالات شدیداً وابسته به یکدیگر از هم جدا شده و در دو گروه متفاوت قرار گیرند، ضریب دو نیمه گاتمن بیشترین مقدار ممکن را خواهد گرفت.

۳. پایایی کادر ریچاردسن (reliability coefficient Kuder- Richardson)

این روش شامل دو آزمون ۲۰KR و ۲۱KR می باشد. این روش ها، آزمون های همگنی (ثبات بین سوالات) می باشند که نسبت پاسخ های صحیح به غلط را در هر سوال ابزار یا آزمون مد نظر قرار می دهند و برای آزمون هایی که پاسخ آن ها به شکل صحیح و غلط می باشد مفید است و در صورتی که کد صفر به پاسخ های غلط و کد یک به پاسخ های صحیح داده می شود به کار رود.

از لحاظ ریاضی قابل اثبات است که ضریب پایایی کادر ریچاردسن برابر با میانگین ضرایب نیمه های آزمون به دست آمده از کل نیمه های ممکن یک آزمون می باشد. در اینگونه آزمون ها (دو حالتی) مقدار ۲۰KR،