

روایی و اعتبار

روایی (validity)

عبارت است از این که یک آزمون تا چه اندازه آنچه را که باید اندازه گیری کند مورد سنجش قرار می دهد.

مثال:

- آیا آزمون بارفیکس قدرت قسمت فوقانی بدن را اندازه گیری می کند؟
- آیا آزمون سنجش چربی زیر پوستی واقعا چربی بدن را اندازه گیری می کند؟

انواع روایی

- روایی منطقی
- روایی آماری

روایی منطقی

این دسته از روایی ها ضعیف ترین انواع روایی هستند زیرا به صورت کیفی و نه کمی بررسی می شوند و از آزمون های آماری استفاده نمی شود.

شامل:

- روایی صوری
- روایی محتوا

روایی صوری

- ضعیف ترین نوع روایی است.
- به نظر بدیهی می رسد که آزمون یا ابزار آنچه را که تصور می شود می خواهد اندازه گیری کند، اندازه گیری نماید.
- مثلاً آزمون دوی ۶۰ متر سرعت دویدن را اندازه گیری می کند.

روایی محتوا

- معمولاً برای آزمون های کتبی در محیط های آموزشی، پرسش نامه ها و دیگر ابزار کتبی زمانی که مقایسه با یک معیار استاندارد ممکن نیست به کار می رود.
- نخستین موضوع در این نوع روایی عبارت است از این که مقولات یا سوالات تا چه اندازه اطلاعات مطلوب را به دقت ارزیابی می کنند.
- معمولاً با جداول مشخص می کنند که چه تعداد از سوالات مربوط به اندازه گیری کدام مقوله است.

روایی آماری

از روایی منطقی قوی تر است چون هم از آزمون های آماری استفاده می کند هم با یک استاندارد مقایسه می شود.

شامل:

• روایی بر مبنی معیار

روایی همزمان

روایی پیش بین

• روایی سازه

روایی بر مبنی معیار

از یک آزمون معیار یا استاندارد طلایی استفاده می شود تا روایی آزمون مورد نظر مشخص شود.

مثال:

- برای اندازه گیری حداکثر اکسیژن مصرفی، دستگاه گاز آنالایزر یک استاندارد طلایی است که می توان برای ارزیابی روایی آزمون های دیگر از آن استفاده نمود.

روایی همزمان

وقتی دو اندازه گیری در شرایط و زمان یکسان انجام شود، یکی از آزمون های بکار رفته آزمون معیار و دیگری آزمونی است که باید روایی سنجی شود.

مثال:

تعیین درصد چربی بدن با روش وزن سنجی زیر آب و روش اندازه گیری چربی زیر پوستی

روایی پیش بین

بررسی ارتباط بین متغیر (های) اندازه گیری شده با وقوع متغیر دیگر

مثال:

محققی آزمون یا پرسشنامه ای را طراحی می کند تا با استفاده از آن مشخص کند که چه بچه هایی احتمالاً ورزشکار نخبه خواهند شد. اگر ارتباط بالایی بین متغیر ها و قهرمان شدن بچه ها در سن بزرگسالی وجود داشته باشد این آزمون دارای روایی بالایی برای پیش بینی نخبه شدن بچه ها می باشد.

روایی سازه

زمانی بکار می رود که هیچ معیار معتبری وجود ندارد و اندازه گیری متغیر مورد نظر دشوار است.

در این مدل افراد به دو گروه (گروهی که درجه بالایی از سازه مورد نظر را دارند و گروهی که درجه پایینی از آن سازه را دارند) تقسیم می شوند. سپس آزمون برای هر دو گروه اجرا می شود و نتایج با آزمون t مقایسه می شود اگر تفاوت قابل ملاحظه ای مشاهده شد یعنی این آزمون روایی بالایی دارد.

مثلا چه عواملی تعیین می کند که افراد به سمت ورزش های رزمی کشیده شوند.

حساسیت (sensitivity)

- درجه ای از دقت آزمون است که تفاوت های بسیار کم را شناسایی می کند.
- مثلاً دقت هر ترازو مشخص است. با باسکول نمی توان طلا را وزن کرد.
- حساسیت کیت های آزمایشگاهی که بر روی بسته آن ذکر شده است در بعضی از تحقیقات ذکر می شود.

اعتبار (پایایی reliability)

سنجش میزان ثبات و قابلیت تکرار نمرات یا داده های آزمون شامل:

- ثبات
- هم ارزی
- ثبات درونی
- پایایی بین آزمون گران یا عینیت

ثبات

در این روش یک آزمون مشابه را در دو موقعیت جداگانه اجرا و نتایج را با هم مقایسه می کنند. میزان همبستگی بین نتایج آزمون ها نشان دهنده پایایی آزمون است.

نکات:

- در این روش اثر یادگیری وجود دارد و لذا برای آزمون های دانشی و مداد و کاغذی استفاده نمی شود.
- زمان بین دو آزمون نباید طولانی باشد چون عواملی مانند رشد یا بلوغ اثر گذار باشند.
- همچنین نباید خیلی کوتاه باشد که خستگی آزمون اول اثر گذار باشد.

هم ارزی

- در این روش دو آزمون متفاوت که متغیر یکسانی را اندازه گیری می کنند اجرا می شوند. سطح همبستگی بین دو آزمون نشان دهنده ثبات دو آزمون است.
- این روش برای متغیرهای دانشی مناسب است البته یکی از ضعف های آن این است که طراحی دو آزمون جامع که آیتم های متفاوت داشته باشند و یک عامل را اندازه گیری کنند، دشوار است.

ثبات درونی

- رایج ترین روش تعیین این نوع پایایی استفاده از دونیم کردن آزمون است (معمولا آیتم های فرد با آیتم های زوج مقایسه می شود) که در آن ارتباط نمرات هر نیمه آزمون با نیمه دیگر بررسی می شود. اگر همبستگی بالا باشد یعنی ثبات درونی بالا است.
- مثلا در یک آزمون تیر اندازی مربی تعداد امتیاز های تیر های زده شده را دو نیم می کند (زوج و فرد) و آنها را با هم مقایسه می کند تا ثبات درونی آزمون خود را محاسبه کند.
- این روش معمولا برای پرسش نامه ها کاربرد دارد که در آن هر چقدر تعداد آیتم ها بیشتر باشد پایایی بدست آمده بیشتر است.

پایایی بین آزمون گران یا عینیت

- در این روش نمرات بدست آمده از آزمون دو آزمونگر با هم مقایسه می شود.
- به عنوان مثال دو محقق درصد چربی را به روش چین پوستی اندازه گیری می کنند.
- این روش برای تحقیق های بزرگ که در آن به تعداد زیادی آزمون گیرنده نیاز است استفاده می شود و باید در تحقیق ذکر شود.
- برای مقایسه باید اندازه گیری های محققین دارای همخوانی بسیار نزدیکی با هم باشند.

ارتباط بین روایی و پایایی

آزمون روا پایا نیز هست اما آزمون پایا ممکن است روایی نداشته باشد.

مثلا تیر انداز را در نظر بگیرید که تمام تیر های او از هدف دور است اما به یک نقطه اصابت کرده است.

خطای اندازه گیری

نمره مشاهده شده = نمره واقعی $\pm$ خطای اندازه گیری

GIGO: آشغال بدهی آشغال می دهد