



روایی و پایایی ابزارهای اندازه‌گیری

Validity & Reliability of measurement tools

Marjan Mirzania

MSc of Health Education & Promotion

Social Development & Health Promotion Research Center

June 28, 2019

مفهوم دقت در اندازه‌گیری

- به عنوان یک اصل کلی هیچ اندازه‌گیری خالی از خطا نیست.
- در یک کلام دقت در اندازه‌گیری یعنی اینکه مقدار سنجیده شده چه میزان به واقعیت نزدیک است.
- دقت در اندازه‌گیری توسط دو شاخص **روایی** و **پایایی** به صورت کمی در می‌آید.

Validity

روایی

- آیا ابزار مورد نظر همان چیزی را می‌سنجد و اندازه‌گیری می‌کند که ما قصد اندازه‌گیری آن را داریم.

Reliability

پایایی

- پایایی یک وسیله اندازه‌گیری به **دقت** آن اشاره دارد.
- یک آزمون در صورتی دارای پایایی است که آن را در یک فاصله زمانی کوتاه چندین بار به گروه واحدی از افراد بدهیم، نمرات حاصل از این چندین بار نزدیک به هم باشند.

Validity

- **Content Validity**
- **Face Validity**
- **Criterion-related Validity**
 - **Predictive V.**
 - **Concurrent V. (Convergent V., Discriminate V.)**
- **Construct Validity**

Validity:

Λ9I!q!fΛ:

1.Content V.

I·COUNfUfΛ:

- هدف

ارزیابی محتوی ابزار و شکل ظاهری آن در اندازه گیری هدف

- سوال

آیا ابزار تهیه شده مفهوم مورد نظر محقق را می سنجد؟

آیا ابزار تهیه شده علاوه بر مفهوم مورد نظر مفاهیم دیگری

را نیز می سنجد؟

آیا ابزار تهیه شده برای سنجش مفهوم کفایت می کند؟

- نحوه ارزیابی

کیفی

کمی

• روش کیفی ارزیابی روایی محتوی

سه مرحله اساسی بررسی کیفی محتوی ابزار توسط پانل:

۱. فیدبک پانل بر اساس معیارهای:

رعایت دستور زبان - Grammar

استفاده از کلمات مناسب - Wording

اهمیت آیتم‌ها - Important

قرارگیری مناسب آیتم‌ها - Item allocation

زمان تکمیل ابزار - Time

امتیازدهی مناسب - Scaling

۲. اصلاح پژوهشگر

۳. تایید اعضای پانل

Validity:

1.Content V.

نحوه ارزیابی:

کیفی
کمی

Validity:

1. Content V.

نحوه ارزیابی:

کیفی

کمی

- روش کمی ارزیابی روایی محتوی:

محاسبه شاخص‌ها

- تعیین نسبت روایی محتوی:

Content Validity Ratio (CVR)

- تعیین شاخص روایی محتوی:

Content Validity Index (CVI)

Validity:

1.Content V.

نحوه ارزیابی:

کیفی

کمی

- روش کمی : محاسبه شاخص‌ها
- تعیین نسبت روایی محتوی (CVR)

- این شاخص توسط لاوشه طراحی شد.
- نخست پانل **ضرورت** هر آیت‌م را بر اساس طیفی سه قسمتی ارزیابی می‌کند:

«ضروری است، E»

«مفید است ولی ضروری نیست، Ubne»

«ضروری نیست، Nn»

- روش کمی : محاسبه شاخص‌ها
- تعیین نسبت روایی محتوی (CVR)

○ سپس نسبت محاسبه می‌شود:

$$CVR = \frac{n_E - \frac{N}{2}}{\frac{N}{2}}$$

Validity:

1.Content V.

نحوه ارزیابی:

کیفی
کمی

Validity:

1. Content V.

نحوه ارزیابی:

کیفی
کمی

- روش کمی : محاسبه شاخص‌ها
- تعیین نسبت روایی محتوی (CVR)

○ مقدار محاسبه شده با مقدار متناظر جدول مقایسه می‌شود:

○ تأیید روایی محتوی **آیتم** اگر:

مقدار محاسبه $<$ مقدار جدول

- روش کمی : محاسبه شاخص‌ها
- تعیین نسبت روایی محتوی (CVR)

Validity:

1.Content V.

نحوه ارزیابی:

کمی، کیفی

جدول تصمیم‌گیری در

مورد CVR

تعداد افراد پانل	حداقل مقدار روایی	تعداد افراد پانل	حداقل مقدار روایی
5	0.99	13	0.54
6	0.99	14	0.51
7	0.99	15	0.49
8	0.85	20	0.42
9	0.78	25	0.37
10	0.62	30	0.33
11	0.59	35	0.31
12	0.56	40	0.29

Validity:

1. Content V.

نحوه ارزیابی:

کیفی
کمی

- روش کمی : محاسبه شاخص‌ها
- تعیین شاخص روایی محتوی (CVI)

- نشان‌دهنده جامعیت قضاوت‌های مربوط به روایی یا قابلیت اجرای مدل، آزمون یا ابزار است.
- جهت بررسی شاخص روایی محتوا از روش والتز و باسل (Waltz & Bausell) استفاده می‌شود.
- پانل هر آیت‌م را بر اساس سه معیار **سادگی**، **اختصاصی/مرتبط بودن**، **وضوح** در طیف لیکرت چهار قسمتی ارزیابی می‌کند:

- روش کمی : محاسبه شاخص‌ها
- تعیین شاخص روایی محتوی (CVI)

Validity:

1.Content V.

○ مثلاً

برای معیار مرتبط بودن:

۱. غیرمرتبط ۲. تا حدودی مرتبط ۳. مرتبط ۴. کاملاً مرتبط

برای معیار ساده بودن:

۱. غیرساده ۲. تا حدودی ساده ۳. ساده ۴. کاملاً ساده

برای معیار واضح بودن:

۱. غیرواضح ۲. تا حدودی واضح ۳. واضح ۴. کاملاً واضح

نحوه ارزیابی:

کیفی
کمی

- روش کمی : محاسبه شاخص‌ها
- تعیین شاخص روایی محتوی (CVI)

Validity:

1. Content V

نحوه ارزیابی:

کیفی

کمی

○ جمع امتیازات موافق در هر آیتم (نمره ۳ و ۴) بر

تعداد کل خبرگان تقسیم می‌شود.

تعداد متخصصینی که به آیتم نمره ۳ و ۴ داده‌اند

CVI= _____

تعداد کل متخصصین

Validity:

1. Content V

نحوه ارزیابی:

کیفی

کمی

- روش کمی : محاسبه شاخص‌ها

- تعیین شاخص روایی محتوی:

Content Validity Index (CVI)

- آیتم‌ها با نمره $CVI > 0.79(0.8)$ پذیرفته می‌شوند.

- CVI برای همه آیتم‌ها محاسبه می‌شود و میانگین آنها به

عنوان CVI مقیاس تعیین می‌شود.

- تأیید روایی محتوی مقیاس اگر:

$CVI > 0.79$

Validity:

1. Content V

Face V.

• هدف

فقط ارزیابی شکل ظاهری ابزار در اندازه‌گیری هدف

• سوال

آیا طراحی ظاهر ابزار برای آزمودنی‌ها و مشاهده‌گران (عوامل اجرایی) مناسب است یا خیر؟

• نحوه ارزیابی

کیفی

کمی

• روش کیفی ارزیابی روایی صوری

سه مرحله اساسی بررسی کیفی ظاهر ابزار توسط خبرگان:

۱. فیدبک خبرگان به لحاظ:

سطح دشواری - Difficulty

میزان عدم تناسب - Irrelevancy

ابهام - Ambiguous

۲. اصلاح پژوهشگر

۳. تایید خبرگان

Validity:

1. Content V

Face V.

نحوه ارزیابی:

کیفی

کمی

• روش کمی ارزیابی روایی صوری

Validity:

1. Content V

Face V.

نحوه ارزیابی:

کیفی

کمی

○ روش تأثیر آیتم: Item Impact Method

○ هدف: کاهش و حذف آیتم‌های نامناسب و تعیین اهمیت هر آیتم

○ نحوه کار: تعیین امتیاز میزان اثر:

$$\text{Impact score} = \text{Frequency (\%)} \times \text{Importance}$$

• روش کمی ارزیابی روایی صوری

○ مثال

اگر امتیازها در پرسشنامه‌ای با مقیاس پاسخ لیکرت ۵ قسمتی:

۱. خیلی ضعیف ۲. ضعیف ۳. متوسط

۴. قوی ۵. خیلی قوی

○ تعیین امتیاز اثر هر آیتم

○ تصمیم‌گیری

اگر امتیاز اثر $< 1/5$ باشد آیتم برای تحلیل‌های بعدی مناسب بوده و حفظ می‌شود.

Validity:

1.Content V

Face V.

نحوه ارزیابی:

کیفی

کمی

Validity:

2. Criterion-related V.

• هدف

تعیین میزان همبستگی / ارتباط بین نمرات حاصل

از یک ابزار اندازه‌گیری (آزمون) با نمرات

حاصل از ابزار دیگر (آزمون ملاک)

Validity:

2. Criterion-related V

2.1. Predictive

2.2. Concurrent

- روایی پیش بین

اگر متغیر ملاک قرار باشد در آینده اتفاق بیافتد و محقق به بررسی ارتباط نتایج حاصل از ابزار خود و متغیر ملاک بپردازد، روایی ملاک از نوع پیش بین یا ناهمزمان است.

- روایی همزمان

اگر متغیر ملاک در همان زمانی که محقق با ابزار خود به تحقیق می پردازد، آماده باشد و محقق بتواند بدون آنکه مدت زمان زیادی را سپری کند به سنجش ارتباط نتایج حاصل از ابزار خود بپردازد، روایی ملاک از نوع همزمان خواهد بود.

Validity:

3. Construct V.

- هدف

بررسی ارتباط درونی متغیرها و مشخص نمودن
متغیرهایی که دارای بیشترین ارتباط

Validity:

3. Construct V.

تعیین روایی سازه

تحلیل عاملی (Factor analysis)

- هدف

- ساده کردن توصیف داده‌ها از طریق تقلیل تعداد متغیرها یا ابعاد مورد مطالعه
- تشخیص عامل‌های مشاهده ناپذیر بر پایه مجموعه‌ای از متغیرهای مشاهده پذیر

Validity:

3. Construct V.

تحليل عاملی:

1. EFA

2. CFA

- تحلیل عاملی اکتشافی:

Exploratory Factor Analysis (EFA)

- تحلیل عاملی تاییدی:

Confirmatory Factor Analysis (CFA)

Validity:

3. Construct V.

تحلیل عاملی:

1. EFA

2. CFA

- وقتی به کار می‌رود که پژوهشگر شواهد کافی قبلی و پیش تجربی برای تشکیل فرضیه درباره تعداد عامل‌های زیر بنایی داده‌ها نداشته باشد.
- تحلیل اکتشافی بیشتر به عنوان یک روش تدوین تئوری و نه یک روش آزمون تئوری در نظر گرفته می‌شود.

Validity:

3. Construct V.

• روش

مولفه‌های اصلی:

Principle Component Analysis (PCA)

تحلیل عاملی:

1. EFA

1.1. PCA

Validity:

- نکته مهم در مورد تحلیل عاملی اکتشافی:

3. Construct V.

پرداختن به نمونه گیری

کیسر مایر الکین (KMO) Kaiser-Meyer-Olkin و

آزمون کرویت بارتلت (BT) Bartlett test و

همچنین مقدار عددی نقطه عطف یا ارزش ویژه (Eigenvalue)

است که معمولا برابر با یک در نظر گرفته می شود.

تحلیل عاملی:

1. EFA

2. CFA

Validity:

3. Construct V.

تحلیل عاملی:

1. EFA

2. CFA

- زمانیکه ابزار یا پرسشنامه موردنظر، از روایی سازه مناسبی برخوردار باشد و جهت تایید کارهای قبلی و بنابراین امکان استخراج فاکتورهای جدید در این نوع تحلیل عاملی وجود ندارد.

Validity:

3. Construct V.

تحليل عاملی:

1. EFA

2. CFA

• نرم افزارها

LISREL ○

AMOS ○

EQS ○

Mplus ○

Smart PLS ○

Validity:

- در خصوص حجم نمونه لازم جهت انجام تحلیل عاملی تناقضات متفاوتی وجود دارد:

3. Construct V.

- Tabachnick & Fidell حداقل ۳۰۰ نمونه را برای انجام تحلیل عاملی ضروری می‌داند.

- برخی منابع دیگر از جمله Lee و Comrey طبقه‌بندی انجام داده‌اند که ۱۰۰ را ضعیف، ۲۰۰ را نسبتاً خوب، ۳۰۰ را خوب، ۵۰۰ را خیلی خوب و ۱۰۰۰ و بیشتر را عالی معرفی کرده‌اند.

تحلیل عاملی:

1. EFA

2. CFA

Reliability

پایایی

- اشاره به دقت یا قابلیت اعتماد یک ابزار (آزمون)
- 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.
- دامنه پایایی از صفر تا یک متغیر است.

Reliability

- **Stability (Consistency)** ○ پایداری (ثبات زمانی)
- **Internal Consistency** ○ همسانی درونی
- **Equivalence** ○ تعادل یا هم‌ارزی

Reliability

- **Stability or Consistency or Stability**
 - Test-retest
- **Internal consistency or Homogeneity**
 - Cronbach's alpha
 - Split half
- **Equivalence**
 - Parallel form

Reliability:

1. Stability

Test-retest

- در این روش آزمون واحدی دوبار روی یک گروه اجرا می‌شود و ضریب همبستگی بین نمرات حاصل از دوبار اجرای آزمون محاسبه می‌شود.

Reliability:

قابل قبولترین روش برای سنجش ثبات:

شاخص همبستگی درون خوشه‌ای

1. Consistency

Interclass Correlation Coefficient (ICC)

Test-retest

○ ثبات مطلوب ابزار اگر:

$$ICC > 0.8$$

Reliability:

2. Internal Consistency

- روش همسانی به جای ثبات نتایج آزمون بر یکنواختی و هماهنگی ماده‌ها یا اجزای تشکیل‌دهنده‌ی یک آزمون تاکید دارد.

Reliability:

2. Internal Consistency

2.1. Cronbach alpha

- در این روش به دلیل این که آزمون فقط یکبار اجرا می شود، اثرات فاصله زمانی بازآزمایی از قبیل حافظه و تمرین به حداقل می رسد.

Reliability:

2. Internal Consistency

2.1. Cronbach alpha

- اگر با حذف سوال میزان آلفا کاهش یابد:
سوال مزبور سهم مثبتی در پایایی آزمون داشته
 - اگر با حذف سوال میزان آلفا افزایش یابد:
سوال دارای همسانی کمتری با سوالات دیگر بوده و
در پایایی سهمی ندارد.
 - تایید پایایی:
- $0.7 > \text{مقدار ضریب آلفای کرونباخ}$

Reliability:

2. Internal Consistency

2.2. Split half

- در این روش آزمون یک بار روی گروهی اجرا و سپس آزمون به دو نیمه تقسیم می شود.

Reliability:

• بهترین روش دو نیمه کردن آزمون:

2. Internal Consistency

2.2. Split half

سوالات فرد در یک گروه و سوالات زوج در گروه دیگر قرار داده شوند. پس برای هر آزمودنی دو نمره (یک نمره برای سوالات زوج و یک نمره برای سوالات فرد) بدست می آید.

سپس ضریب همبستگی حاصل از نمرات دو نیمه آزمون محاسبه خواهد شد که فرمول پیشنهادی برای محاسبه ضریب همبستگی نمرات دو نیمه آزمون، فرمول اسپیرمن-براون است.

Reliability:

2. Internal Consistency

2.3. Kuder-Richardson

• کاربرد

برای آزمون‌هایی که سوالات آنها با هم همگن نیستند یا سوال‌های آن از گروه‌های مختلف تشکیل شده‌اند و یا نمره‌گذاری سوالات با هم یکسان نیست.

- روش کودر-ریچاردسون زمانی به کار می‌رود که سوالات به صورت صفر و یک نمره‌گذاری می‌شود.

Reliability:

3. Equivalence

- ناظر بر وضعیتی است که طی آن نتایج حاصل از ابزاری که محقق برای سنجش متغیر مورد نظر خود تهیه نموده است با نتایج حاصل از ابزار دیگر که برای سنجش همان متغیر تهیه شده و قبلاً روایی و پایایی آن اثبات شده است شباهت داشته باشد.

Reliability:

- در این روش از یک آزمون دو فرم موازی یا هم ارز تهیه می کنند.

3. Equivalence

3.1. Paraller form

- دو فرم آزمون همزمان یا با فاصله ی زمانی کوتاهی روی یک گروه واحد اجرا می کنند و ضریب همبستگی بین نمرات حاصل از دو فرم آزمون را محاسبه می کنند.

عوامل موثر بر روایی آزمون

- کیفیت سؤالهای آزمون
- ترتیب قرار گرفتن سؤالهای آزمون
- درجه دشواری سؤالهای آزمون
- طول یک آزمون یا تعداد سؤال ها
- اجرای نامناسب آزمون
- اشکالات موجود در تصحیح پاسخ های آزمون شوندگان و نمره گذاری
- ترکیب آزمون شوندگان از لحاظ توانایی مورد اندازه گیری
- ویژگیهای روانی آزمون شوندگان

عوامل موثر بر پایایی آزمون

- با افزودن بر تعداد سؤالات یک آزمون، پایایی آن نیز افزایش می‌یابد.
- با متجانس‌تر و همگون‌تر کردن سؤالات یک آزمون، پایایی آن بیشتر می‌شود.
- با افزودن تعدادی سؤال با ضریب تمیز زیاد به یک آزمون، پایایی آن بالا می‌رود.
- با افزودن تعدادی سؤال دارای ضریب دشواری متوسط به یک آزمون، پایایی آن افزایش می‌یابد.
- اگر یک آزمون را در مورد گروهی از افراد که از نظر توانایی مورد سنجش تجانس کمتری با یکدیگر دارند اجرا کنیم، آزمون پایایی بیشتری نشان خواهد داد.
- آزمون‌های سرعت معمولاً از آزمون‌هایی که سرعت عملکرد در آنها عامل مهمی به حساب نمی‌آید (آزمون‌های قدرت) ضریب پایایی بیشتری را نشان می‌دهند.

رابطه بین روایی و پایایی از این قرار است:

- روایی **شرط کافی** برای پایایی است.

پس هرچه معتبرتر باشد پایا نیز است.

- پایایی **شرط لازم** برای روایی است.

پس یک اندازه پایا ممکن است معتبر نباشد.

هر روایی می تواند پایا باشد

اما

هر پایایی نمی تواند روا باشد.

- حیدری چروده م. راهنمای سنجش روایی و پایایی در پژوهش‌های فرهنگی و اجتماعی. مشهد، جهاد دانشگاهی مشهد، ۱۳۸۹.
- محمدیگی ا، محمدصالحی ن، گل م. روایی و پایایی ابزارها و روش‌های مختلف اندازه‌گیری آنها در پژوهش‌های کاربردی در سلامت. مجله دانشگاه علوم پزشکی رفسنجان، ۱۳۹۳.
- گنجی ح، ثابت م. روان‌سنجی (مبانی نظری آزمون‌های روانی). موسسه نشر ساوالان، ۱۳۷۹.
- حاجی زاده، ابراهیم. اصغری، محمد. روش‌ها و تحلیل‌های آماری با نگاه به روش تحقیق در علوم زیستی و بهداشتی. جهاد دانشگاهی. چاپ اول ۱۳۹۰.
- Di lorio CK. Reliability assessment and item analysis, Validity. Measurement in Health Behavior: Methods for Research and Education. Jossey-Bass, 2005; 163-236.
- Munro B. H. Statistical methods for healthcare research & application of SPSS in data analysis. 2004.
- Lawshe CH. A quantitative approach to content validity. Personnel Psychology, 1975; 28(4): 563-575.
- Tabachnick BG, Fidell LS. Using multivariate statistics. 5th ed. Boston: Allyn and Bacon; 2007.

A bouquet of pink daisies with yellow centers and long green palm fronds is arranged in the background. The text is overlaid on the flowers.

Thank You

Have a great day!