

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

کارگاه روایی و پایایی در اندازه‌گیری علوم پزشکی

الهه طالبی قانع

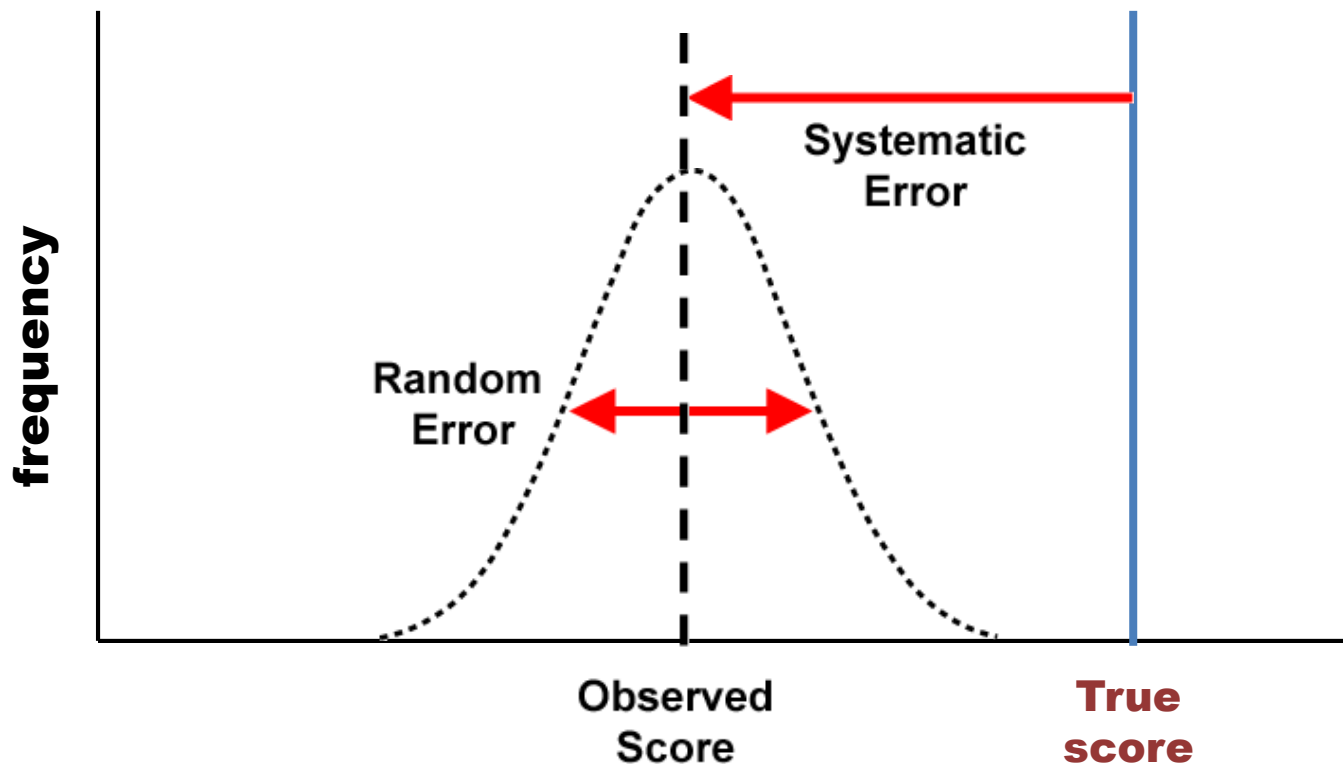
talebi_ghane@yahoo.com

- بخش اول) مقدمه
 - مفاهیم و تعاریف
 - انواع خطاها
 - تعریف دقت و صحت
 - آشنایی با روایی و پایایی و ارتباط بین آنها
- بخش دوم) روایی در اندازه‌گیری علوم پزشکی
 - آشنایی با انواع روایی
 - محاسبه روایی با استفاده از نرم‌افزار Excel
- بخش سوم) پایایی در اندازه‌گیری علوم پزشکی
 - آشنایی با انواع پایایی
 - محاسبه پایایی با استفاده از نرم‌افزار SPSS

بخش اول: مقدمه

مفاهیم و تعاریف

- داده‌ها از طریق اندازه‌گیری به دست می‌آیند.
- اندازه‌گیری تخصیص اعداد به افراد، رویدادها یا مشاهدات براساس قوانینی خاص است که می‌تواند به صورت خودگزارشی، مشاهده، پرسشنامه، مصاحبه، تجهیزات علمی (ترازو، دستگاه فشارخون)، ابزار پزشکی، تجهیزات آزمایشگاهی و ... جمع‌آوری گردد.
- اهمیت جمع‌آوری داده‌هایی با کیفیت بالا
- عدم جمع‌آوری داده‌هایی با کیفیت بالا منجر به دو نوع خطا می‌شود:
 - خطای سیستماتیک
 - خطای تصادفی



- انحراف از مقدار واقعی اندازه (خطای سیستماتیک) ← عدم صحت
- مقداری از تغییرپذیری ذاتی در ابزار اندازه‌گیری (خطای تصادفی) ← عدم دقت

چنین خطاهای اندازه‌گیری می‌تواند بصورت جدی بر تفسیر اندازه‌ها بدون توجه به اینکه آنها از چه ابزاری (اشعه ایکس، تست‌های آزمایشگاهی یا مصاحبه) بدست آمده‌اند، تاثیر بگذارد.

۱- خطای سیستماتیک: انحراف از مقدار واقعی اندازه است.

- عدم صحت-عدم روایی

۲- خطای تصادفی: مقداری از تغییرپذیری ذاتی در ابزار اندازه گیری است.

- عدم دقت - عدم پایایی

Validity and Reliability

روایی؟؟؟

پایایی؟؟؟

صحت و اعتبار (Validity)

- توافق بین مقادیر اندازه‌گیری شده و مقدار واقعی که بوسیله یک دستگاه/ابزار بدون خطا بدست آمده است.
- اندازه‌گیری واقعی متغیرها
- از طرف دیگر به قدرت تعمیم‌پذیری نتایج مطالعه به جامعه بزرگتر و به موقعیت‌های واقعی اطلاق می‌شود.
- ▶ آزمونی دارای روایی است که برای اندازه‌گیری آنچه مورد نظر است مناسب باشد.

دقت و پایایی (Reliability)

- دقت یک ابزار اندازه‌گیری به معنای توانایی آن ابزار برای تولید دوباره همان نتایج است. داشتن دقت منجر به بالارفتن پایایی (Reliability) می‌شود.
- چنانچه متغیر مورد سنجش دو یا چند بار با همان وسیله (یا وسیله مشابه و قابل مقایسه با آن) و در شرایط مشابه اندازه‌گیری شود، نتایج مشابه و یکسانی به دست آید که نشان دهنده ثبات و استمرار اطلاعات جمع‌آوری شده در دفعات مختلف با استفاده از آن وسیله اندازه‌گیری و قابلیت تکرارپذیری و تولید مجدد همان اطلاعات است.

ارتباط بین روایی و پایایی

- برای داشتن یک اندازه‌گیری صحیح هر وسیله اندازه‌گیری باید پایا و معتبر باشد.
- یک آزمون باید پایا باشد تا بتواند روا باشد یعنی پایایی شرط روایی است اما روایی برای پایایی ضروری نیست.
- از طرف دیگر یک وسیله اندازه‌گیری حتماً باید معتبر باشد تا پایا بودن آن ارزشمند باشد.

Validity and Reliability



Reliable
Not Valid



Low Validity
Low Reliability



Not Reliable
Not Valid



Both Reliable
and Valid



بخش دوم: روایی در اندازه گیری های پزشکی

تعریف روایی

- آیا ابزار مورد نظر می تواند به قدر کافی **معنای واقعی** مفهوم مورد نظر را منعکس کند؟
- ابزاری دارای روایی است که برای اندازه گیری آنچه که مورد نظر است **کافی و مناسب** باشد

روایی:

میزان توانایی آزمون در اندازه گیری آن چیزی که باید اندازه بگیرد

انواع روایی

- **روایی ترجمه (Translation Validity)**
- **روایی وابسته به محتوا (Content-related Validity)**
 - **Content Validity** روایی محتوا
 - **Face Validity** روایی ظاهری
- **روایی وابسته به ملاک (Criterion-related Validity)**
 - **Predictive**
 - **Concurrent**
- **روایی وابسته به سازه (Construct-related Validity)**
 - **Factor Analysis**
 - **Convergent**
 - **Divergent**
 - **Discrimination**

Translation Validity

روایی ترجمه فرایندی است برای بالا بردن روایی ظاهری و محتوایی یک پرسشنامه ی ترجمه شده.

شیوه ی Forward-backward:

۱. ترجمه ی پرسشنامه از زبان اصلی به زبان فارسی توسط چند متخصص (هم از لحاظ زبان و هم رشته ی مربوطه)
۲. تشکیل یک گروه از افراد متخصص و انتخاب یک نسخه فارسی واحد
۳. بازگردانی نسخه فارسی به زبان اصلی توسط چند متخصص (ترجیحا دو زبانه)
۴. تشکیل یک گروه از افراد متخصص و انتخاب یک نسخه باز ترجمه شده ی واحد
۵. نسخه باز ترجمه شده به همراه نسخه اصلی برای طراح پرسشنامه فرستاده شود و توسط وی ارزیابی گردد.
۶. انجام سایر مراحل اعتبار سنجی (بررسی اعتبار)

Content-related Validity

روایی وابسته به محتوا

Content-related validity

روایی محتوی نیاز به داوری و قضاوت متخصصان و صاحب نظران

اظهار نظر در مورد اینکه:

آیا خصوصیت مورد نظر را اندازه می گیرد؟

آیا گویا، روشن، بدون ابهام و با واژه مناسب است؟

آیا مقیاس مناسبی دارد؟

آیا وزن ارائه شده برای حیطة ها و گزینه ها مناسب است؟

آیا متناسب با شخصیت آزمون شوندگان است؟

آیا زیر مجموعه های (حیطه های) ارائه شده مناسبت دارند؟

آیا گزینه ها معرف مجموعه ها هستند؟

آیا توالی و انسجام خوبی دارد؟

Content-related validity

روایی ظاهری

روایی محتوا

- روایی محتوا (content validity) یک پرسشنامه به این معناست که سوال های آن، جنبه های مختلف متغیر مورد سنجش را به میزان کافی و مناسب تحت پوشش قرار دهد.

Content-related validity

شاخص پیشنهادی لاوشه (Lawshe) برای ارزیابی روایی محتوا:

Content Validity Ratio (CVR)

برای هر سوال در پرسشنامه محاسبه می شود

ابتدا از تعدادی متخصص درخواست می شود که بیان کنند حضور هر یک از سوالات در پرسشنامه، **اساسی**، **مفید** و یا **غیر ضروری** است.

Content-related validity

$$CVR = \frac{n_e - n/2}{n/2}$$

که در آن،

n_e : تعداد متخصصانی که سوال را “اساسی” دانسته اند

n : تعداد کل متخصصان


$$-1 \leq CVR \leq 1$$

Number of panelists	Min acceptable CVR value
5	0.99
6	0.99
7	0.99
8	0.78
9	0.75
10	0.62
11	0.59
12	0.56
13	0.54
14	0.51
15	0.49
20	0.42
25	0.37
30	0.33
35	0.31
40	0.29

Content-related validity

- اگر CVR بزرگتر یا مساوی حداقل مقدار قابل پذیرش باشد، سوال (گویه) پذیرفته می شود
- اگر CVR کوچکتر از صفر باشد، سوال حذف می شود
- در غیر این صورت به شکل زیر عمل می کنیم:
به گزینه **اساسی** نمره ی ۳، به **ضروری** نمره ۲ و **غیر ضروری** نمره ۱ می دهیم و برای هر سوال میانگین نمرات داوران را حساب می کنیم
اگر بیشتر یا مساوی ۲ باشد ← سوال پذیرفته می شود
در غیر این صورت ← رد می شود.

Content-related validity

در نهایت با میانگین گرفتن از CVR مربوط به سوالات پذیرفته شده مقدار میانگین CVR برای کل پرسشنامه به دست می آید.
یک شاخص برای کل پرسشنامه:

مقدار آن باید بیشتر از $0/7$ باشد.

Content-related validity

Content Validity Index (CVI)

روش دیگر برای اطمینان از روایی محتوی، تایید سوالات پرسشنامه از جنبه مرتبط بودن، واضح بودن و ساده بودن، توسط چند نفر از متخصصین و آزمون شوندگان است

مرتبط بودن: (نا مرتبط است (۱)، تا حدودی مرتبط است (۲)، بسیار مرتبط است (۳)، کاملاً مرتبط است (۴)

واضح بودن: (نا واضح است، نیاز به اصلاح اساسی دارد، واضح است اما نیاز به اصلاح جزئی دارد، کاملاً واضح است)

ساده بودن: (ساده نیست، نیاز به اصلاح اساسی دارد، ساده است اما نیاز به اصلاح جزئی دارد، کاملاً ساده است)

Content-related validity

در اینصورت شاخص روایی محتوا (CVI) به صورت زیر قابل محاسبه است:

$$CVI = \frac{\# 3 \ \& \ 4}{\# Total}$$

سوالاتی که CVI آنها بالاتر از ۰/۷۵ باشد مناسب هستند.

توجه: ابتدا مربوط بودن سوالات را می سنجیم، سپس دو مورد دیگر را بررسی می کنیم.

Content-related validity

روایی ظاهری (face validity) بیان کننده آن است که تا چه حد، مخاطب از سوالات همان چیزی را که مقصود طراح پرسشنامه بوده است، برداشت می کند.



Content-related validity

هر چند که روایی ظاهری یک ویژگی مطلوب آزمون است، با این حال در بعضی آزمون ها این ویژگی نه تنها ضروری نیست بلکه بهتر است که وجود نداشته باشد، به عنوان مثال اگر کسی بخواهد بیماری **روانی** را در افراد تشخیص دهد، بهتر است آزمونی را به کار بندد که روایی ظاهری **اندکی** دارد.

Criterion-related Validity

روایی وابسته به ملاک

Criterion-related Validity

- منظور از روایی ملاک ،میزان ارتباط بین نمرات حاصل از **یک آزمون** با نمرات حاصل از **یک آزمون** یا وسیله اندازه گیری **دیگر** است که به صورت **gold standard** باشد.
- بنابراین هر زمان که از نمرات یک آزمون (پیش بین) برای پیش بینی عملکرد افراد (در آینده یا در زمان حال) در آزمونی دیگر(آزمون ملاک) استفاده کنیم با روایی ملاک سر و کار داریم.

Predictive Validity “پیش بین”

Concurrent Validity “همزمان”

روایی ملاک :

Criterion-related Validity

“پیش بین” Predictive Validity

در صورتی آزمون، روایی پیش بین دارد که بتواند عملکرد خاصی را در آینده پیش بینی کند

توانایی پیش بینی چیزی بر اساس یک ویژگی

مثال: آزمون استعداد تحصیلی (The Scholastic Aptitude Test (SAT)) یک پیش بین معتبر برای متمایز ساختن دانش آموزانی که در آینده در دانشگاه عملکرد مناسبی دارند با آنهایی که عملکرد مناسب ندارند، می باشد.

Criterion-related Validity

- اغلب برای غربال کردن (انتخاب نامزد مناسب) به کار می رود. (مثال پیشگیری از طلاق)
- بررسی آن نیاز به گذشت زمان دارد
- از معایب آن زمان بر بودن و هزینه بر بودن است.

Criterion-related Validity

“همزمان” Concurrent Validity

روایی هم زمان، توانایی آزمون را در تخمین عملکرد خاصی در زمان حال نشان می دهد

آیا یک آزمون یا وسیله اندازه گیری می تواند جانشین مناسبی برای آزمون یا وسیله اندازه گیری دیگری (بدون فاصله و بطور همزمان) باشد؟

- همبستگی بالا
- مقرون به صرفه بودن اقتصادی / زمان / آسان بودن

Criterion-related Validity

روش های بررسی روایی ملاک:

- ضریب همبستگی (بالای ۰/۷)
- آزمون مقایسه میانگین ها
- حساسیت و ویژگی

زمانی به کار می رود که آزمون ملاک (یا gold standard) وجود داشته باشد

Criterion-related Validity

مثال: بررسی تست اثر کف در شناسایی موارد لوله گذاری داخل تراشه مشکل در بیماران دیابتی بستری در بیمارستان های وابسته به دانشگاه علوم پزشکی شهید بهشتی

جمع	لوله گذاری مشکل		تست اثر کف
	دارد	ندارد	
۴۷	۱۱	۳۶	آسان) $3 >$
۲۸	۱۹	۹	سخت) $3 \leq$
۷۵	۳۰	۴۵	جمع

Criterion-related Validity

حساسیت، ویژگی، ارزش اخباری مثبت و ارزش اخباری منفی دو تست اثر کف در بیماران دیابتی کاندید اعمال جراحی انتخابی در بیمارستان‌های منتخب وابسته به دانشگاه شهید بهشتی در سال ۱۳۹۳.

تست اثر کف دست	تست
	پارامترها
۶۳/۳	حساسیت
۸۰	ویژگی
۶۷/۹	ارزش اخباری مثبت
۷۶/۶	ارزش اخباری منفی

Construct-related Validity

روایی وابسته به سازه

Construct-related Validity

سازه عبارت است از یک صفت، ویژگی یا کیفیت روان شناختی (چیزی که نمی توان آن را مستقیماً مشاهده نمود) که از طریق نظریه های روان شناسی قابل استنتاج است.
مثال: هوش، خلاقیت، عشق و ...

اعتبار سازه ای یعنی میزان توانایی پرسشنامه در اندازه گیری سازه یا صفت به خصوصی از فرد

روایی سازه بر این مطلب اشاره دارد که آیا ابزار اندازه گیری مورد نظر با مفاهیم نظری هم خوانی دارد یا نه؟

Construct-related Validity

جنبه های مختلف تعیین روایی سازه:

- **روایی همگرا (Convergent Validity)**
- **روایی واگرا (Divergent Validity)**
- **روایی افتراقی (Discriminant Validity)**
- **تحلیل عاملی (Factor Analysis)**
- **همسانی درونی (Internal Consistency)**

Construct-related Validity

روایی همگرا (Convergent Validity)

این روایی به میزان ارتباط بین اندازه‌ی موجود با اندازه‌هایی که از لحاظ نظری انتظار می‌رود با هم مرتبط باشند اطلاق می‌شود.

یک ابزار با متغیرهایی که از
لحاظ تئوری انتظار می‌رود
مرتبط باشند، همبستگی بالایی
داشته باشد

مانند: هوش و پیشرفت تحصیلی

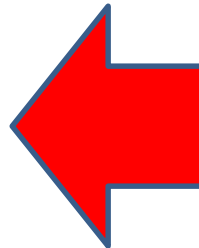
این روایی را می‌توان با استفاده از معیار همبستگی بررسی کرد

Construct-related Validity

معمولا برای تایید روایی سازه از شواهد آماری همراه با استدلال منطقی استفاده می شود.

برای تایید روایی سازه، نیاز است که روش های آماری، نظریه ای که آزمون بر آن استوار است را تایید نمایند

اندازه گیری هوش با این ابزار
باید بتواند بین هوش و
سرعت یادگیری ارتباط
معناداری بیابد و نشان دهد
که پیشرفت تحصیلی با میزان
هوش رابطه مستقیم دارد



سازه هوش:
• افراد با هوش یادگیری سریع
تری دارند
• هوش با پیشرفت تحصیلی
رابطه مستقیم دارد

Construct-related Validity

Table 2. Correlation matrix for different domains of WHOQOL-100 and WHOQOL-BREF (n=60)

Domain	WHOQOL-100						
	Physical	Psychological	Level of Ind.	Social relation.	Environment	Spirituality /religion.	
WHOQOL-BREF							
Physical health	r*	0.759	0.268	0.694	0.465	0.358	0.456
	P**	<0.001	0.037	<0.001	<0.001	0.005	<0.001
Psychological health	r	0.354	0.739	0.603	0.589	0.550	0.671
	P	0.006	<0.001	<0.001	<0.001	<0.001	<0.001
Social relationships	r	0.216	0.195	0.427	0.665	0.541	0.484
	P	0.098	0.132	0.001	<0.001	<0.001	<0.001
Environmental health	r	0.081	0.017	0.297	0.297	0.683	0.413
	P	0.537	0.896	0.020	0.020	<0.001	0.001

* Pearson's correlation coefficient

** Two-tailed P-value

Construct-related Validity

روایی واگرا (Divergent Validity)

آزمون مورد نظر با آزمون های دیگری که سازه های متفاوتی را اندازه گیری می کنند، همبستگی نداشته باشد یا نتایج معکوس داشته باشد.

سازه های مربوط به پرسشنامه اضطراب با سازه های مربوط به نشاط زندگی

Construct-related Validity

روایی افتراقی (Discriminant Validity)

ابزار یا آزمون مورد نظر قادر باشد گروه های متفاوت از هم را تمیز دهد.

مثلا پرسشنامه کیفیت زندگی نمرات متفاوتی را در بیماران و افراد سالم گزارش دهد.

Construct-related Validity

Table 6. The comparison of different facets and domains of WHOQOL-100 between healthy and disabled participants

	Healthy		Disabled		P*	P**
	Mean	SD	Mean	SD		
Domain						
Physical	13.53	2.25	12.50	2.15	<0.001	<0.001
Psychological	12.46	1.08	12.25	1.02	<0.001	<0.001
Level of independence	14.06	1.76	13.18	2.29	<0.001	<0.001
Social relationship	12.98	1.83	12.68	1.19	<0.001	<0.001
Environment	12.70	1.76	12.30	1.86	<0.001	<0.001
Spirituality/religion/personal beliefs	14.65	3.39	14.66	3.50	0.888	0.915

* P-value based on independent samples t-test

** P-value based on multiple linear regression analysis

Construct-related Validity

مثال:

آزمون تفکر انتقادی باید بتواند بین دانشجویان فلسفه و پرستاری اختلاف معناداری به دست آورد.

Construct-related Validity

تحلیل عاملی (Factor Analysis)

روش آماری برای تعیین تعداد و ماهیت متغیرهایی که یک خصوصیت را مشخص می کند

** اگر یک پرسشنامه برای اندازه گیری سازه ی کیفیت زندگی است، ترکیب عاملی آن باید نشان دهنده عوامل مطرح در سازه کیفیت زندگی باشد.

Construct-related Validity

تحليل عاملی می تواند به دو صورت انجام گیرد:

- تحليل عاملی اکتشافی (Exploratory Factor Analysis (EFA))

- تحليل عاملی تاییدی (Confirmatory Factor Analysis (CFA))

Construct-related Validity

همسانی درونی (Internal Consistency)

- وجود همبستگی بین نمرات زیر گروه ها (حیطه ها) با نمره کل آزمون
- نبود همبستگی زیاد بین گروه ها (حیطه ها) - زیرا در غیر این صورت همه ی آنها یک چیز را اندازه گیری می کنند نه حیطه های مختلف را.

با تشکر از توجه شما