

سمینار و روش تحقیق اعتبارسنجی

دکتر مهدی روانشادنیا

فهرست بخش هفتم: اعتبارسنجی

مفاهیم پایه اعتبارسنجی

انواع اعتبار

عوامل موثر بر پایایی و اعتبار

شیوه های اعتبارسنجی

قضاوت آماری

مفاهیم پایه اعتبارسنجی

مفهوم شناسی
(اعتبار)
Validity

مفهوم اعتبار به این سؤال پاسخ می‌دهد که ابزار اندازه‌گیری تا چه حد خصیصه مورد نظر را می‌سنجد.

نکته

بدون آگاهی از اعتبار ابزار اندازه‌گیری نمی‌توان به دقت داده‌های حاصل از آن اطمینان داشت.

مفهوم دقت در اندازه گیری

➤ به عنوان یک اصل کلی هیچ اندازه گیری خالی از خطا نیست.

➤ در یک کلام دقت در اندازه گیری یعنی اینکه مقدار سنجیده شده چه میزان به واقعیت نزدیک است.

➤ دقت در اندازه گیری توسط دو شاخص اعتبار و پایایی به صورت کمی در می آید.

پایایی یا ثبات (reliability)

در صورت تکرار یک اندازه گیری تا چه میزان مقادیر سنجیده شده به یکدیگر نزدیک هستند.

اگر دو مرتبه از یک فرد به فاصله چند دقیقه و در شرایط کاملاً یکسان فشارخون گرفته شود تا چه میزان مقادیر به یکدیگر نزدیک هستند؟

اگر به فاصله چند روز از یک فرد در خصوص تعداد ساعات مشاهده تلویزیون سوال شود آیا اعداد بیان شده به یکدیگر نزدیک هستند؟

اعتبار (validity)

- مقدار سنجیده شده چه میزان به واقعیت نزدیک است؟
- اندازه گیری فشارخون با یک کاف کوچک هرچند ممکن است پایا باشد ولی معتبر نیست.
- سوال در خصوص بعضی موضوعات خاص ممکن است پایا باشد ولی معتبر نباشد.
- میزان درآمد شما در ماه چقدر است؟
- در هفته چند ساعت کتاب می خوانید؟
- چه میزان سیگار کشیدن را مضر می دانید؟ (به خصوص اگر از افراد سیگاری پرسیده شود)
- چند مرتبه فرزندتان را کتک زده اید؟

رابطه بین پایایی و اعتبار

➤ پایایی شرط لازم برای اعتبار است.

➤ پس یک اندازه پایا ممکن است معتبر نباشد.

➤ اعتبار شرط کافی برای پایایی است.

➤ پس هرچه معتبر است پایا نیز است.

Verification vs validation

► **Verification:**

"Are we building the product right"

- The software should conform to its specification

► **Validation:**

"Are we building the right product"

- The software should do what the user really requires

Verification vs Validation

► **Verification:**

"Are we building the product right"

► The software should conform to its specification

► **Validation:**

"Are we building the right product"

► The software should do what the user really requires

انواع اعتبار

روش‌های تعیین اعتبار اندازه‌گیری

۱. اعتبار محتوا: ملاک بررسی اجزای تشکیل دهنده یک ابزار اندازه‌گیری است.
۲. اعتبار ملاکی: کارآمدی یک ابزار اندازه‌گیری در پیش بینی رفتار یک فرد در موقعیت‌های خاص
۳. اعتبار سازه: ابزار اندازه‌گیری تا چه اندازه یک سازه را می‌سنجد.
۴. اعتبار عاملی: از طریق تحلیل عاملی بدست می‌آید.

انواع اعتبار ملاکی

بررسی رابطه نمره‌های آزمودنی که برخی ویژگی‌ها را می‌سنجد و آنچه ادعای پیش‌بینی آن را دارد.

اعتبار
پیش‌بین

مواردی که داده‌های حاصله از دو اندازه‌گیری در یک زمان در دسترس باشد.

اعتبار
همزمان

شیوه های تعیین اعتبار سازه

تفاوت های گروهی

تغییرات

(تشخیصی) اعتبار همگرا و واگرا

وجوه مختلف اعتبار نه انواع آن

اعتبار ظاهری یا صوری [face validity]

آیا پاسخ دهنده مفهوم متن را همان می یابد که محقق مد نظر داشته است؟

آیا ظاهر متن به اندازه کافی شکیل است که حواس پاسخ دهنده را جلب کند؟

اعتبار محتوا [content validity]


محتوای پرسشنامه در راستای هدف اصلی و سوال پژوهش است؟

اعتبار ساختاری [construct validity]

اسکلت پرسشنامه با مفهوم انتزاعی موضوع تطابق دارد؟

اعتبار پیشگویی کننده [predictive validity]

پاسخهای داده شده تا چه میزان با نتایج آینده همخوانی دارد؟



اعتبار داخلی: روایی پایایی

اعتبار داخلی:

روایی:

آیا وسیله ای که برای اندازه گیری متغیر خاصی در نظر گرفته شده می تواند به درستی آن متغیر را اندازه گیری کرده و وضعیت آن را مشخص نماید؟

تعیین روایی:

• **اعتبار محتوا** **Content validity**: یک پارچگی و درهم تنیدگی ابزار را تعیین می کند (قضاوت افراد متخصص)

• **اعتبار ملاکی** **Criterion validity**: همبستگی نمره ابزار با نمره ملاک (رغبت شغلی در ۱۸ سالگی با ملاک در ۲۸ سالگی)

• **اعتبار سازه** **Construct validity**: تطابق نظریه با نتیجه ابزار (تفاوت بلوغ اجتماعی در کودکان و بزرگ سالان)



اعتبار داخلی پایایی Reliability :

یک وسیله در پروژه و در شرایط یکسان تا
چه اندازه ای نتایج یکسانی به دست می
دهد؟

اعتبار داخلی پایایی Reliability :

۱- ثبات **stability**: اعلام یک نتیجه در اندازه گیری های مکرر (ضریب همبستگی)

۲- سنخیت **Homogeneity**: یک ابزار فقط یک موضوع را سنجش کند (افسردگی، وسواس) {سازگاری داخلی Int.consistency }

۴- هم قدر بودن **Equivalence**: توافق مشاهدگرها



اعتبار خارجی : گستره تعمیم پذیری یافته ها

شیوه نمونه گیری
تعداد نمونه

عوامل موثر بر پایایی و اعتبار

عوامل موثر بر پایایی و اعتبار

- شیوه طراحی سوالات
- نوع سوالات و گزینه ها
- ترتیب سوالات
- جامع و مانع
- فرد پرسشگر
- شرایط پرسشگری در یک محیط متناسب و آرام
- شکل ظاهری پرسشنامه

عوامل موثر بر اعتبار و پایایی (۱)

➤ شیوه طراحی سوالات:

➤ کدام سوال احتمالاً معتبرتر است؟

➤ در سال گذشته چند بار سردرد داشته اید؟

➤ در ماه گذشته چند بار سردرد داشته اید؟

➤ در دو روز گذشته چند بار سردرد داشته اید؟

➤ اگر بخواهیم در خصوص بستری در بیمارستان سوال کنیم فکر می کنید بهترین بازه زمانی چقدر است؟

عوامل موثر بر اعتبار و پایایی (۲)

نوع سوالات و گزینه ها

بعضی مواقع گذاردن گزینه باعث سوگیری پاسخها می شود. لذا در این مواقع سوالات کوتاه پاسخ ارجح هستند.

گزینه ها گاه با مفهوم سوال همخوانی ندارد.

چه اندازه رفتار پرستاران با بیماران مناسب است؟

بسیار خوب، خوب، نمی دانم، بد، بسیار بد

بسیار خوب، خوب، تا حدودی خوب، بد، بسیار بد

بسیار خوب، خوب، نیمه مطلوب، بد، بسیار بد

صحیح، تا حدودی درست، میانه، بد، غیر مسولانه

راضی، تا حدودی راضی، کمی راضی، ناراضی، بسیار ناراضی

موافقم، تا حدودی موافقم، بی نظرم، مخالفم، بسیار مخالفم

همیشه، اغلب اوقات، گاهی اوقات، بسیار کم، ابداً

۰، ۲۵٪، ۵۰٪، ۷۵٪، ۱۰۰٪

عوامل موثر بر اعتبار و پایایی (۳)

ترتیب سوالات

سوال در مورد خصوصیات و مسایل حساس در ابتدای پرسشنامه اعتبار سوالات بعدی را ممکن است به شدت کاهش دهد.

به هم ریختگی زیاد و ارتباط مبهم سوالات ممکن است باعث مخشوش شدن ذهن پاسخ دهنده شود.

گذاردن سوالات انگیزشی در ابتدای پرسشنامه و همچنین مقدمه روشن و کامل در ابتدای پرسشنامه و یا قبل از هر سوال، اعتبار سوالات را افزایش می دهد.

عوامل موثر بر اعتبار و پایایی (۴)

جامع و مانع و مختصر و مفید باشد

مثال: کدام سوال معتبرتر است؟

نظرتان در خصوص رفتار دانشجویان با بیماران چیست؟

دانشجویان بایست بتوانند در طول دوره خود شیوه صحیح ارتباط برقرار نمودن با بیمار را فراگیرند و به شکلی منطقی با بیمار ارتباط برقرار کنند. ارتباط بسیار نزدیک با بیمار و یا ارتباطی یک طرفه و آمرانه مطلوب نیست و باعث می شود که بیمار به خوبی به پزشک خود اعتماد نکند. با توجه به این توضیحات فکر می کنید شیوه ارتباط دانشجویان با بیماران چگونه است؟

عوامل موثر بر اعتبار و پایایی (۵)

➤ فرد پرسشگر

➤ شکل ظاهری، موقعیت اجتماعی، اخلاق و رفتار

➤ آیا پاسخهای دریافت شده توسط پرسشگران زن و مرد برابر است؟

➤ آیا پاسخ داده شده توسط بیمار به استاد معتبرتر است یا به دانشجو؟

➤ آیا پاسخ بیمار به دانشجویی که مرتب است و خود را معرفی می کند دقیقتر است

یا به دانشجویی که با لباس چروکیده و با چهره ای نامرتب دارد؟

➤ به نظر شما سن پرسشگر و سواد وی در دریافت پاسخها تاثیر دارد؟

عوامل موثر بر اعتبار و پایایی (۶)

➤ شرایط پرسشگری در یک محیط متناسب و آرام (اگر بتوان پیدا نمود!!!)

➤ سوال در خصوص کیفیت تدریس در زمان امتحان

➤ پرسشگری از مردم در هنگام رانندگی و پشت چراغ قرمز

➤ هنگام گردش در پارک

عوامل موثر بر اعتبار و پایایی (۷)

➤ شکل ظاهری پرسشنامه

➤ تعداد صفحات و تعداد سوالات

➤ رنگ کاغذ

➤ فونت مورد استفاده

➤ آرایش متن

➤ بکار بردن کلمات ملموس و مورد استفاده در فرهنگ مربوطه

➤ دقت در عدم استفاده از کلمات منفی به صورت مکرر در جمله

شیوه های اعتبارسنجی

نحوه سنجش پایایی

- **Temporal stability**
 - Retest with a reasonable gap
- **Form equivalence**
 - Parallel tests
- **Internal consistency**
 - the coefficient of test scores obtained from a single test or survey (Cronbach Alpha, KR20, Spilt-half).

شیوه سنجش اعتبار پرسشنامه (۱)

- باید دقت شود که اعتبار پرسشنامه با اعتبار پاسخها یکی نیست.
- برای سنجش اعتبار باید gold standard داشت که در مورد پرسشگری تقریباً در اکثر موارد وجود ندارد.
- آیا نظرات کارشناسان را می توان ملاک قرار داد و به عنوان نوعی gold standard دانست؟
- در صورت مثبت بودن پاسخ شاید یک راه درخواست از کارشناسان برای نظر دادن به کیفیت تک تک سوالات و یا گروه سوالات است.

شیوه سنجش اعتبار پرسشنامه (۲)

➤ به این شکل که حدود ۵ تا ۱۰ نفر کارشناس به کیفیت سوالات امتیاز داده و سپس امتیازات آنها با یکدیگر جمع زده و میانگین گیری می شود.

به عنوان یک اصل شخصاً هیچ یک از ضرایب محاسبه شده به این شیوه را باور نمی کنم مگر خلاف آن ثابت شود. اگرچه متأسفانه خود مجبورم که در مشاوره ها دانشجویان را مجبور به انجام اشتباهی بکنم که خود به آن اعتقاد ندارم.

➤ به نظر من نظرات تنها یک کارشناس دقیق و کارپشته بسیار با ارزشتر از جمع نظرات تعداد زیادی کارشناس است که یا دقیق جواب ندهند و یا با موضع آشنا نباشند.

شیوه سنجش اعتبار پاسخها

- مصاحبه با پاسخ دهندگان و تعیین میزان دقت پاسخ مکتوب آنها
- سنجش متغیرها پرسشنامه با شیوه عینی در صورت امکان
- مقایسه پاسخها با نتایج gold standard در صورت وجود
- گذارن check questions و حذف پاسخهای غیر دقیق و همچنین تعیین میزان خطاهای موجود در پاسخها به عنوان شاخصی از اعتبار پاسخها

قضاوت آماری

حجم منطقی نمونه مورد مطالعه

اگر هدف از تحقیق برآورد فراوانی یک نسبت است به عبارتی مقیاس سنجش متغیر مورد نظر کیفی باشد می توان از فرمول زیر برای محاسبه تعداد نمونه استفاده کرد:

$$n = \frac{z^2 p(1 - p)}{d^2}$$

$$n = \frac{z^2 \times s^2}{d^2}$$

اگر هدف از تحقیق برآورد میانگین یک متغیر است به عبارتی مقیاس سنجش متغیر مورد مطالعه کمی باشد برای محاسبه تعداد نمونه فرمول زیر بکار برده می شود:



95%	1.95
97.5%	2.24
99%	2.57

روش نمونه گیری

۱- سرشماری (Census) که در آن کلیه افراد جامعه مورد نظر تحت بررسی قرار می گیرند.

۲- نمونه گیری (Sampling) که در آن تنها بخشی از افراد جامعه مورد بررسی قرار می گیرند. این بخش از جامعه باید طوری انتخاب شود که کلیه خصوصیات مهم جامعه را داشته باشد (معرف جامعه باشد) تا نتایج حاصل از این نمونه قابل تعمیم به آن جامعه باشد.

روشهای متداول انتخاب نمونه بر دو گونه است:

۱- نمونه گیری احتمالی (Probability sampling)
که در آن هر فرد از جامعه دارای شانس معین (و غیر صفر !
(جهت انتخاب شدن در نمونه است.

۲- نمونه گیری غیر احتمالی (Non-probability
Sampling) که انتخاب افراد در این روش بر اساس
تصادف و شانس نبوده و بیشتر بعلت محدودیت امکانات تیم
تحقیق بکار گرفته می شود.

نمونه گیری احتمالی:

- ۱- نمونه گیری تصادفی ساده (Simple random sampling)
- ۲- نمونه گیری با طبقه بندی (Stratified sampling)
- ۳- نمونه گیری خوشه ای (Cluster sampling)
- ۴- نمونه گیری منظم (Systematic sampling)
- ۵- نمونه گیری چند مرحله ای (Multistage sampling)

نمونه گیری غیراحتمالی:

- ۱- نمونه گیری آماده یا در دسترس (Convenience sampling)
- ۲- نمونه سهمیه ای (Quota sampling)

قضاوت آماری

« تعیین رابطه سطح سواد والدین با وضعیت امتحان نهایی کودکان »

معدل	ضعیف	متوسط	خوب	عالی	جمع
					سطح سواد والدین
بالا	۱۰	۲۰	۳۰	۴۰	۱۰۰
پائین	۵۰	۸۰	۲۰	۵۰	۲۰۰

فرضیه :

معدل امتحان نهایی کودکان با والدین دارای سطح سواد بالا و پایین یکسان است.

معدل امتحان نهایی کودکان با والدین دارای سطح سواد بالا و پایین فرق می کند.

قضاوت آماری

« تعیین رابطه سطح سواد والدین با وضعیت امتحان نهایی کودکان »

Chi-square test گروه های مستقل و متغیر وابسته کیفی است

قضاوت آماری

« تعیین رابطه سطح سواد والدین با وضعیت امتحان نهایی کودکان »

Chi-square test

- مشخص کردن اینکه آیا بین گروه‌ها از نظر موارد مشاهده شده (Observed) اختلافی وجود دارد یا نه
- محاسبه مقادیر مورد انتظار (Expected) مربوط به هر خانه (مقداری که اگر رابطه‌ای وجود نداشته باشد حاصل می‌شود) = (جمع هر ستون × جمع هر سطر) تقسیم بر جمع کل
- محاسبه عدد Chi-square هر خانه
- ۴ - تعیین عدد Chi-square کلی
- ۵ - تعیین P value

قضاوت آماری

« تعیین رابطه سطح سواد والدین با وضعیت امتحان نهایی کودکان »

Chi-square test

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

قضاوت آماری

« تعیین رابطه سطح سواد والدین با وضعیت امتحان نهایی کودکان »

Chi-square test

$P < 0.0005$

قضاوت آماری

« تعیین رابطه سطح سواد والدین با وضعیت امتحان نهایی کودکان »

Chi-square test

$P < 0.0005$

تحقیق نشان داد که کودکان دارای والدین با سطح سواد بالا نسبت به کودکان دارای والدین با سطح Chi-square سواد پائین از نظر پراکندگی به لحاظ وضعیت امتحان نهایی تفاوت دارند و با آزمون $(p < 0.0005)$ مشخص شد که چنین اختلافی از نظر آماری معنی دار است.

قضاوت آماری

تعیین اثر Propranolol بر پرفشاری متوسط خون

۲۰ نفر مبتلا به پرفشاری متوسط خون با در نظر گرفتن اصول حاکم بر این گونه تحقیقات انتخاب و به ۱۱ نفر آنان Propranolol و به ۹ نفر دارونما (Placebo) داده شد. در خاتمه تحقیق فشار خون افراد هر دو گروه اندازه گیری شد.

قضاوت آماری

تعیین اثر Propranolol بر پرفشاری متوسط خون

فرضیه تحقیق:

اثر Propranolol و دارونما بر فشار خون مبتلایان به پرفشاری متوسط خون یکسان است.

اثر Propranolol و دارونما بر فشار خون مبتلایان به پرفشاری متوسط خون یکسان نیست.

فشار خون بیماران دریافت کننده Propranolol در مقایسه با دریافت کنندگان دارو نما کاهش

می یابد.

فشار خون بیماران دریافت کننده Propranolol در مقایسه با دریافت کنندگان دارو نما افزایش

می یابد.

قضاوت آماری

تعیین اثر **Propranolol** بر پرفشاری متوسط خون

Student t test دو گروه مستقل و متغیر وابسته کمی است

قضاوت آماری

بر پرفشاری متوسط خون Propranolol تعیین اثر

Student t test دو گروه مستقل و متغیر وابسته کمی است

مراحل انجام **Student t test** :

- محاسبه میانگین متغیر وابسته در هر گروه
- محاسبه انحراف معیار متغیر وابسته در هر گروه
- محاسبه عدد **t**
- تعیین **P value**

قضاوت آماری

تعیین اثر Propranolol بر پرفشاری متوسط خون

Student t test دو گروه مستقل و متغیر وابسته کمی است

$$sd = \sqrt{\frac{\sum (X - \bar{X})^2}{n - 1}}$$
$$sd = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}}$$

$$t = \frac{|x_1 - x_2|}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)} \times \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

بر پرفشاری متوسط خون Propranolol تعیین اثر

گروه گواه	گروه تجربی		گروه گواه	گروه تجربی	
	مجدور هر انحراف از میانگین	مجدور هر انحراف از میانگین		مجدور هر انحراف از میانگین	مجدور هر انحراف از میانگین
۶۴	-۸	۱۳۲	۱۳۲	-۱۳	۱۰۷
۱۴۰	+۲۰	۱۴۰	۱۴۰	-۱۸	۱۰۲
۰	۰	۱۴۰	۱۴۰	+۵	۱۲۵
۶۴	-۸	۱۳۲	۱۳۲	+۹	۱۲۹
۰	۰	۱۴۰	۱۴۰	+۸	۱۲۸
۱۹۶	+۱۴	۱۵۴	۱۵۴	-۶	۱۱۴
۱۶	-۴	۱۳۶	۱۳۶	-۵	۱۱۵
۳۶	-۶	۱۳۴	۱۳۴	+۶	۱۲۶
۶۴	-۸	۱۳۲	۱۳۲	+۳	۱۲۳
				+۳	۱۲۳
				+۸	۱۲۸
۸۴۰	صفر	۱۲۶۰	۱۲۶۰	صفر	۱
<u>میانگین=۱۴۰</u>		<u>میانگین=۱۲۰</u>			

$$140-120=20$$

$$(20*100) : 140 = 14.3\%$$

شماره	گروه تجربی		گروه گواه		
	انحراف هر یک از میانگین	مجدور هر انحراف از میانگین	فشار خون سیستولی	انحراف هر یک از میانگین	مجدور هر انحراف از میانگین
۱۰۷	-۱۳	۱۶۹	۱۳۲	-۸	۶۴
۱۰۲	-۱۸	۳۲۴	۱۶۰	+۲۰	۴۰۰
۱۲۵	+۵	۲۵	۱۴۰	۰	۰
۱۲۹	+۹	۸۱	۱۳۲	-۸	۶۴
۱۲۸	+۸	۶۴	۱۴۰	۰	۰
۱۱۴	-۶	۳۶	۱۵۴	+۱۴	۱۹۶
۱۱۵	-۵	۲۵	۱۳۶	-۴	۱۶
۱۲۶	+۶	۳۶	۱۳۴	-۶	۳۶
۱۲۳	+۳	۹	۱۳۲	-۸	۶۴
۱۲۳	+۳	۹			
۱۲۸	+۸	۶۴			
۱	صفر	۸۴۲	۱۲۶۰	صفر	۸۴۰
۳					
۲۰					
	<u>میانگین=۱۲۰</u>		<u>میانگین=۱۴۰</u>		

$$140-120=20$$

$$(20*100) : 140=14.3\%$$

تحقیق نشان داد که **Proranolol** در مقایسه با دارو نما فشار خون را در مبتلایان به پرفشاری متوسط خون ۲۰ میلیمتر جیوه (**14.3%**) کاهش می دهد .

تحقیق نشان داد که **Proranolol** در مقایسه با دارو نما فشار خون را در مبتلایان به پرفشاری متوسط خون ۲۰ میلیمتر جیوه (**14.3%**) کاهش می دهد .

$P < 0.001$

تحقیق نشان داد که **Propranolol** در مقایسه با دارونما فشار خون را در مبتلایان به پرفشاری متوسط خون ۲۰ میلیمتر جیوه **14.3 %** کاهش می دهد. چنین کاهش به لحاظ آماری معنی دار است (**$P < 0.001$**).

تکالیف بخش اعتبارسنجی پژوهش

- روشهای اعتبارسنجی مدل‌های فراابتکاری را شناسایی و معرفی کنید.
- برای قضاوت‌های آماری نمونه‌هایی از صنعت ساخت بیاورید.