

НАДЕЖНОСТЬ - показатель качества метода и данных, получаемых с помощью этого метода, мера независимости данных от случайных обстоятельств и нерелевантных факторов.

Классическая формулировка Н. В классической теории измерения Н. метода понимается как его устойчивость к меняющимся условиям применения, в том числе нечувствительность к влиянию исследователя, или как согласованность измерения. Результат измерения включает истинный компонент и ошибку. Последняя представляет собой суперпозицию влияний, не коррелирующих между собой.

Это предположение делается для того, чтобы гарантировать нормальное распределение ошибок и предсказывать их поведение. Классические показатели Н., используемые для оценки качества опросников, концептуально определяются как дополнение до единицы отношения дисперсии ошибки к дисперсии измерения. Традиционно обозначается r_{xx} : это корреляция измеренной переменной самой с собой, очищенная от ошибки измерения. Другая интерпретация показателя Н. - квадрат корреляции между истинными и измеренными значениями. Истинное значение понимается как сумма ответов респондента на все возможные вопросы, измеряющие данный конструкт, напр., социально-экономический статус. Отсюда следует, что чем большую долю универсума вопросов (индикаторов) включает методика, тем выше ее Н.

Одним из методов эмпирической оценки Н. является повторное применение методики на одних и тех же респондентах (ретестовая Н.). Корреляция между результатами двух измерений интерпретируется как коэффициент Н. Этот подход основывается на нереалистичном предположении о стабильности истинного показателя во времени и не принимает в расчет реактивный эффект (изменение свойств респондента вследствие первого измерения, в особенности при обсуждении сензитивных тем). В результате корреляция между двумя измерениями ниже действительной Н. методики.

Эффекты памяти, напротив, завышают показатель Н. Предположение об отсутствии корреляции между ошибками двух измерений также нереалистично. Ретестовый подход к определению Н. сложен технически: исследователь должен идентифицировать индивидуальные результаты обоих измерений и обеспечить повторное участие респондентов в процедуре.

Метод альтернативных форм основан на использовании двух взаимозаменяемых методик. В некоторых случаях исследователям доступны альтернативные шкалы аттитюдов (см.), или ценностей. Возражением против этого подхода является неправдоподобность предположения о полной эквивалентности разных версий одной методики.

Метод расщепления методики пополам строится на предположении о строгой параллельности двух частей методики. Корреляция половин между собой является оценкой H . каждой половины. H . методики находится по формуле

Спирмена-Брауна. Этот подход не предполагает устойчивости истинного показателя во времени, исключает эффект памяти и возможность ковариации ошибок измерения. Однако полученный показатель H . всегда зависит от способа деления методики на половины. Обобщением этого подхода является метод внутренней согласованности.

В некоторых случаях и сама измеряемая область поведения, и выборка вопросов считается однородной, а все элементы методики - коррелирующими. H . методики определяется по формуле Л. Кронбаха:

где k - количество элементов измерительной методики, σ^2_t - дисперсия суммарных измеренных баллов, σ^2_i - дисперсия отдельных элементов. Формула имеет интересную теоретическую интерпретацию и более широкий диапазон пригодности. Она находит применение для разработки шкал Лайкерта. Близкий подход связан с подстановкой вместо дроби в правой части формулы Кронбаха величины, обратной первому собственному значению из анализа главных компонент индикаторов.

Более современный и общий подход к определению H . связан с использованием структурных уравнений (см.). Последний метод позволяет, в частности, оценивать H . менее "родственных" измерений - т. наз. тау-эквивалентных и параллельных.

H . как воспроизводимость результатов, полученных разными исполнителями. В исследованиях, связанных с использованием наблюдения или кодирования текстового материала в качестве меры H . используется сходство результатов, полученных разными исполнителями ("экспертами"), работающими по одной схеме (инструкции).

Результаты измерения (кодирования) сводятся в таблицу сопряженности, строки и столбцы которой состоят из категорий, использованных, соответственно, первым и вторым исполнителем. Коэффициент согласованности k ("каппа") Коэна рассчитывается по формуле:

$$k = (D - P) / (1 - P) ,$$

где $D = (1/N) \sum_{i=1}^I X_{ii}$, т. е. сумма частот по диагонали, деленная на общее количество наблюдений (чистая согласованность), а $P = (1/N^2) \sum_{i=1}^I X_{i\cdot} X_{\cdot i}$, т. е. сумма попарных произведений сумм по столбцу и по строке с одним и тем же номером, деленная на квадрат общего количества наблюдений (случайная согласованность). Похожий подход

используется для проверки правильности ("объективности") транскрипции устных текстов в нарративном анализе, анализе разговора и т.п.

В случае, когда измерение осуществляется несколькими экспертами, используется коэффициент конкордации Кендалла.

Процедурные средства обеспечения Н. В тех случаях, когда Н. методики не может быть оценена количественно, исследователи предпринимают усилия для того, чтобы идентифицировать и устранить источники случайной изменчивости данных. Этнографы используют стандартизованные процедуры наблюдения, множественные источники фиксации материала, стремятся к максимально полной и точной фиксации происходящего, приглашают независимых наблюдателей.

Этнометодологи и разговорные аналитики используют лонгитюдный дизайн исследования, чтобы избежать случайной фиксации нетипичных и мимолетных событий и случайной потери долговременных закономерностей. Исследователи текста сравнивают результаты, полученные при анализе различных фрагментов целостного текста. Визуальные социологи используют несколько параллельно средств фиксации происходящего в разных ракурсах. Важным принципом оценки Н. по-прежнему остается воспроизводимость данных/результатов несколькими независимыми учеными. Многие методологи качественных исследований неоднократно пытались отказаться от понятия объективности и Н., считая его мерой нечувствительности метода к социальным изменениям и контексту. С этой точки зрения независимость данных, полученных в интервьюировании, объявлялась фикцией и артефактом исследования.

Однако даже те, кто понимает интервьюирование как совместное творчество, используют повторяемости и регулярности текста как подтверждение достоверности данных. Во многих отраслях качественных исследований используются процедуры оценки полученных материалов, напоминающие финансовый аудит (см. Валидность).

С.В. Сивуха