信度是测量结果的稳定性程度。

如果能用统一测量工具反复测量某人的同一种心理特质，则其多次测量的结果间的一致性程度就叫做信度，有时也叫做测量的可靠性。

## 信度定义：

1. 被测团体的真分数变异数与实得分数变异数之比。
2. 一个被试团体的真分数与实得分数的相关系数的平方。

以上两条，只具有理论意义

1. 一个测验X与他人以一个平行测验X’的相关系数。

第三条具有实际意义。

描述测量一致性程度的指标还可以用**信度指数**，实际上是信度系数的平方根。

## 信度的作用：

1. 信度是测量过程中所存在的随机误差大小的反应。

测量中的系统误差与信度无关。因为系统误差只对测量结果产生恒定的影响，不会产生上下波动。

如果信度很低，随机误差就很大，测量的结果就会与真分数产生较大偏差。

1. 信度可以用来解释个人测验分数的意义。

我们可以用一个团体的两次施测结果代替对一个人反复施测（来获得真分数），以估计测量误差的变异数。这是，每个人两次测量的分数之差可以构成一个新的分布，这个分布的标准差就是测量的标准误。

一个测量的标准误

* SEM=st(1-rtt)1/2
  + SEM=标准误
  + st =测验分数的标准差
  + rtt =测验的信度

经典测量理论三大假设

* + 1 对一个人测量无数次，其平均误差为0，即E=0
  + 2真分数与误差没有关系，即rTE=0
  + 3 不同测量上的误差没有关系，即rE1E2=0
  + 如果进行无数次测量，X的平均值等于T

当测验满足经典测量理论的三大假设是，就可以用以下方式构建测验真分数估计的置信区间

x-ZcSe<=T<=x+ZcSe

Se是标准误 Zc是对应于某个统计检验显著水平标准正态分布下的临界值

1. 信度可以帮助进行不同测验分数的比较

差异分数的信度估计(Gulliksen, 1950)

rdiff=(1/2(rxx+ryy)-rxy)/(1-rxy)

(rxx和ryy=测验X、Y的信度

rxy=测验X和Y的相关

* **差异分数的信度远低于各测验的信度，是测验间相关的函数**

## 信度估计的方法

一 重测信度

用同一个量表对同一组被试施测两次所得结果的一致性程度，其大小等于同一组被试在两次测验上所得分数的皮尔逊积差相关系数。

只能在允许重测的情况下进行计算。具体说，他必须满足三个条件

1. 所测量的心理特性必须是稳定的
2. 遗忘和练习的效果基本上相互抵消
3. 在两次施测的间隔期内，被试在所要测查的心理特质方面没有获得更多的学习和训练。

报告重测信度时，应说明两次施测的间隔以及被试在此期间内的有关经历。

二 复本信度

同一批被试在两个复本测验上所得分数的皮尔逊积差相关系数

如果两个复本测验连续施测，则称为等值性系数。它主要反映两个复本测验题目差别所带来的变异情况。

为抵消施测的顺序效应，一般可以采用平衡设计。

报告时 也应说明两次施测的间隔以及被试在此期间内的有关经历。

能一定程度克服重测信度练习效应的影响

三 分半信度

指将一个测验分成两半后，所有被试在这两半上所得分数的一致性程度。

解释与等值性系数一样，可以看成在最短时距内施测的两个平行测验，

有时被称作内部一致性系数。

用等值复本信度计算方法计算后，还要用斯皮尔曼布朗公式加以校正

* rxx=n\*rhh/(1+(n-1)\* rhh)
  + n=改变后的测验与原测验长度之比
  + 在这里，n=2
  + rhh=校正前的测验信度
  + rxx =校正后的测验信度
* 例子：100个题目，分半信度r=.70,整个测验的估计信度是多少？
* rxx=(2\*.7)/1+(2-1)\*.7)=1.4/1.7=.82

也就是 2Rhh/(1+Rhh)

斯皮尔曼布朗公式只有在两半测验分数的方差相等时才能使用。

（否则应该使用弗朗那根公式或者卢纶公式）

分半信度通常是在只能施测一次或者没有复本的情况下使用。

同一个测验会有很多个分半信度值。

四 内部一致性系数（同质性信度）

测验内部所有题目间的一致性程度。

包含两种含义：

1. 所有题目都是测得同一种心理特质
2. 所有题目得分之间都具有较高的正相关

当一个测验具有较好的同质性信度时，说明测验主要测的是某一个单个心理特质，实测结果就是该特质水平的反映。

库—里信度系数

* 库德-理查德森公式(Kuder-Richardson formula 20, 21)
* rxx=k[1-Σpi(1-pi)/s2]/(k-1) K-R20
* rxx= [ k-x(k-x)/ s2 ]/(k-1) K-R21
  + - k=测验的题目数
    - s2 =整个测验得分的方差（变异数）
    - pi =被试作对题目I的比例
* K-R21假设所有题目具有相等的难度，更保守的估计
* 例：如果一个测验有75题，平均值=50，方差=100，则
  + rxx = [75-50(75-50)/100]/74=.84
* **注意**：K-R20,K-R21只适用与0-1记分的测验

科隆巴赫系数

* Alpha系数(Cronbach,1951)
* 概念上，Alpha是经斯皮尔曼-布郎公式校正过的所有分半信度的平均值。它是一个一般公式。
* 统计上
* Alpha=(k2\*r)/R 或
* Alpha=k(1- Σsi2/st2)/(k-1)
  + - k=测验的题目数
    - r =题目间的相关之平均，不包括相关矩阵中对角线上的值
    - R=相关矩阵中所有题目间相关的总和
    - si2 =题目I得分的方差（变异数）
    - st2 =整个测验得分的方差（变异数）

不同测验的信度要求（Alpha系数）

|  |  |  |
| --- | --- | --- |
| **信度估计** | **测验类型** | **解释** |
| .95 |  | 测量误差确实没有影响 |
| .90 | 标准化的群体智力测验 |  |
| .85 | 标准化的成就测验 | 中到高的信度 |
| .80 | 自陈报告式人格测验 |  |
| .75 | 课堂上多项选择测验 | 低到中的信度 |
| .70 | 等级量表 |  |
| .65 |  | 低信度 |
| .60 | 某些投射测验 |  |
| .55 |  |  |
| .50 |  | 真分数和误差分数对测验分数有同等影响 |

五 评分者信度

多个评分者给同一批人的答卷进行评分的一致性程度

评分者人数为两个时，可采用相关系数（积差相关或者等级相关）

评分者人数多于两个时，可采用肯德尔和谐系数进行估计。

肯德尔和谐系数参见心统ppt“相关”部分

六 合成分数的信度

* 本质上等同于信度的原始定义。
* k-(k\*rii)
* Rss=1- -------------
* k+(k2-k)\*rij
* rss = 总的信度
* rii = 测验的平均信度
* K = 测验数量
* rij = 测验间的平均相关
* K2=测验数量的平方

七 相关的校正

* 我们更想估计两个构想的真正相关，而不是由于不可信测量而得到的低估了的相关。
* 为此，我们可以运用校正公式估计在度量是完美的情况下两个构想的相关性（理论关系）
* rc=rxy/(rxx\*ryy)1/2
* rxy=X与Y的实得相关
* rxx与ryy=X 与Y的信度系数

## 影响测量信度的主要因素

1 被试方面

就个体而言，身心健康状况、应试动机、注意力、耐心、求胜心、作答态度都会影响测量误差

就团体而言，整个团体内部水平的离散程度以及团体平均水平会影响信度。（因为我们计算信度估计值大都是以相关为基础的，而相关系数的大小往往取决于全体被试得分的分布情况。当背时团体内部差异很大时，全体被试的总分分布会较广，以相关为基础计算出来的信度值会偏大，会高估信度值。）

2 主试者方面

施测者遵守指导手册的规定的程度

评分者评分标准掌握情况

3 施测情境

4 测量工具

5 两次施测间隔时间

我的疑惑求大牛解答：

1 ppt55页高保真是神马玩意儿，，，

2

Version1.0 Made by 严嘉