

心理测量第四讲

讲师 司马紫衣



高途学院APP下载



高途学院公众号





第八章 测验等值

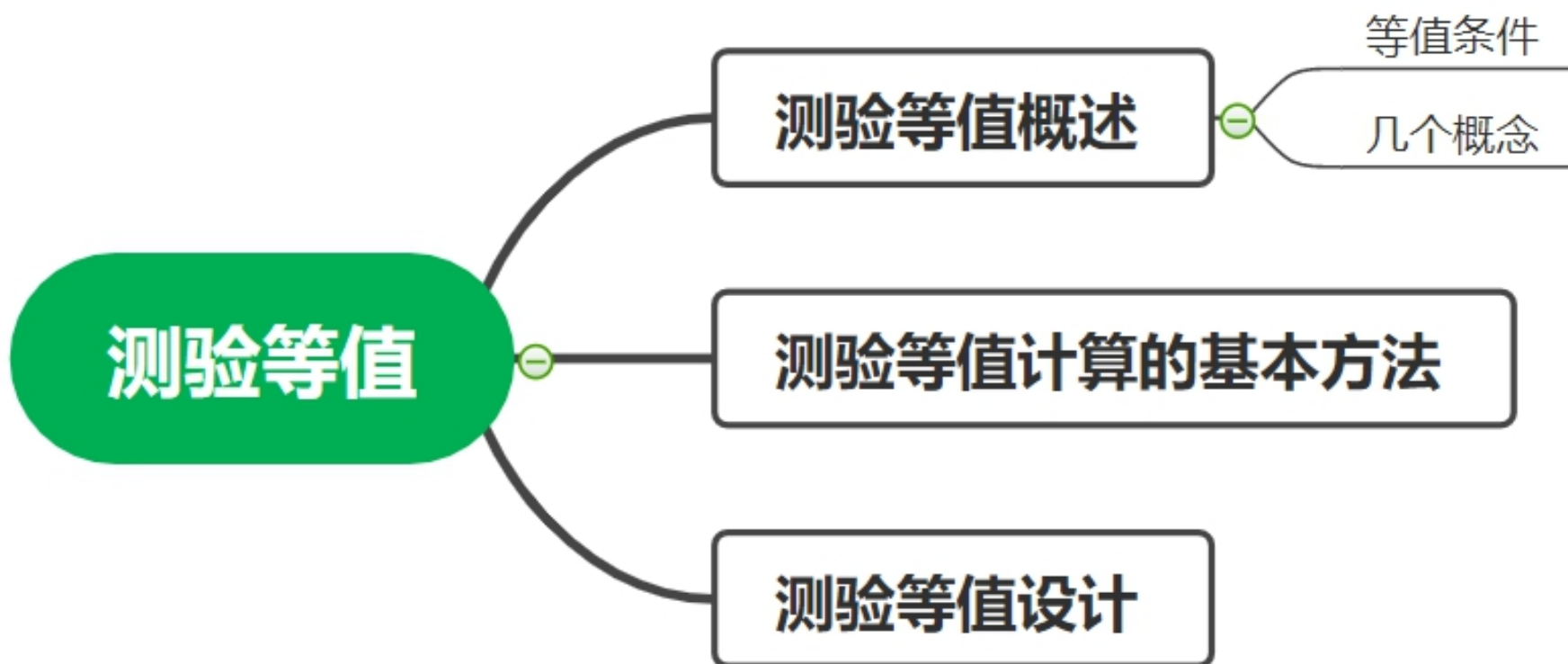


概览

- ◆ 测验等值概述（选择）
- ◆ 测验等值计算的基本方法（选择）
- ◆ 常用测验等值设计介绍（选择）



本章结构





一、测验等值概述

(一) 测验等值的实质

- 通过对考核同一种心理品质的多个测验形式做出测量分数系统的转换，进而使这些不同测验形式的测验分数之间具有可比性。



一、测验等值概述

(二) 测验等值的条件

1. **同质性**：测同一种心理品质，内容与范围基本相同。
2. **等信度**：相等的测验信度。
3. **公平性**：考生参加被等值的任何一个测试，等值后的结果都是一样的。
4. **可递推性**：测验X与测验Y等值，测验Y与测验Z等值，可以递推出测验X与测验Z之间的等值关系。
5. **对称性**：从等值的两个测验中的任何一个测验出发，所获得等值对应关系是相同的。
6. **样本不变性**：等值关系不因测试样本的变化和测验情境的变化而发生变化。



一、测验等值概述

(三) 测验等值的一些基本概念

1. 测验分数等值：直接操作对象是**原始分数**。
2. 项目参数等值：直接操作对象是**项目参数**。
3. 水平等值：两测验**难度水平大体相当**，受测团体能力分布相似。
4. 垂直等值：两测验**难度水平有明显差异**，受测团体能力水平不相同。
5. 等值标准误差：由**抽样引起的等值误差**称为等值的**随机误差**，评价等值随机误差大小的指标称为等值标准误差。
6. 等值偏差：**等值处理方法不当**引起的等值误差，称为等值的**系统误差**，也称等值偏差。



一、测验等值概述

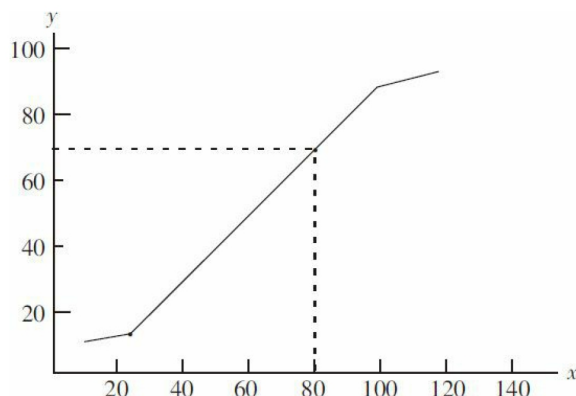
(四) 测验等值结果的表示方法

1. 表列法

x	45	50	55	60	65	70	75	80	85
y	38.27	42.52	47.11	51.70	56.30	61.26	66.62	72.15	78.35
PR	6.225 2	12.384 1	20.562 9	31.092 7	44.205 3	59.006 6	72.615 9	83.344 4	92.019 9

2. 公式法: $y = Ax + B$

3. 图示法





二、测验等值关系计算的基本方法

(一) 等百分位等值

- 两个分数，一个在测验形式 x 上，另一个在测验形式 y 上，如果这两个分数在各自测验中所处的百分等级相等，这两个分数就被认为是等值的。



二、测验等值关系计算的基本方法

(二) 线性等值

- 两个分数，一个在测验形式 x 上，另一个在测验形式 y 上，如果这两个分数在各自测验中的标准分数相等，这两个分数就被认为是等值的。



三、常用测验等值设计介绍

- 随机分组——每组实施一个测验
- 随机分组——各测验对每组都实施
 - ✓ 施测顺序相反
- 随机分组——每组各实施一个测验，锚测验向每组实施
- 非随机分组——每组各实施一个测验，锚测验向每组实施



真题演练

如果要建立两个测验的等值关系，这两个测验须满足的前提条件是（ ）。（统考，2013）

- A.信度相等
- B.长度相等
- C.效度相等
- D.测验总分的标准差相等

根据经典测量理论，常用的测验等值方法有（ ）。（多选）（统考，2008）

- A.全距等值
- B.等百分等值
- C.线性等值
- D.方差等值



真题演练

如果要建立两个测验的等值关系，这两个测验须满足的前提条件是（ ）。 （统考，2013）

- A.信度相等
- B.长度相等
- C.效度相等
- D.测验总分的标准差相等

答案：A

根据经典测量理论，常用的测验等值方法有（ ）。 （多选） （统考，2008）

- A.全距等值
- B.等百分等值
- C.线性等值
- D.方差等值



真题演练

如果要建立两个测验的等值关系，这两个测验须满足的前提条件是（ ）。 （统考，2013）

- A.信度相等
- B.长度相等
- C.效度相等
- D.测验总分的标准差相等

答案：A

根据经典测量理论，常用的测验等值方法有（ ）。 （多选） （统考，2008）

- A.全距等值
- B.等百分等值
- C.线性等值
- D.方差等值

答案：BC



真题演练

表示测验等值结果的方法有（ ）。（多选）（统考，2018）

A.图示法

B.公式法

C.列表法

D.数轴法



真题演练

表示测验等值结果的方法有（ ）。（多选）（统考，2018）

A.图示法

B.公式法

C.列表法

D.数轴法

答案：ABC



第九章 目标参照测验

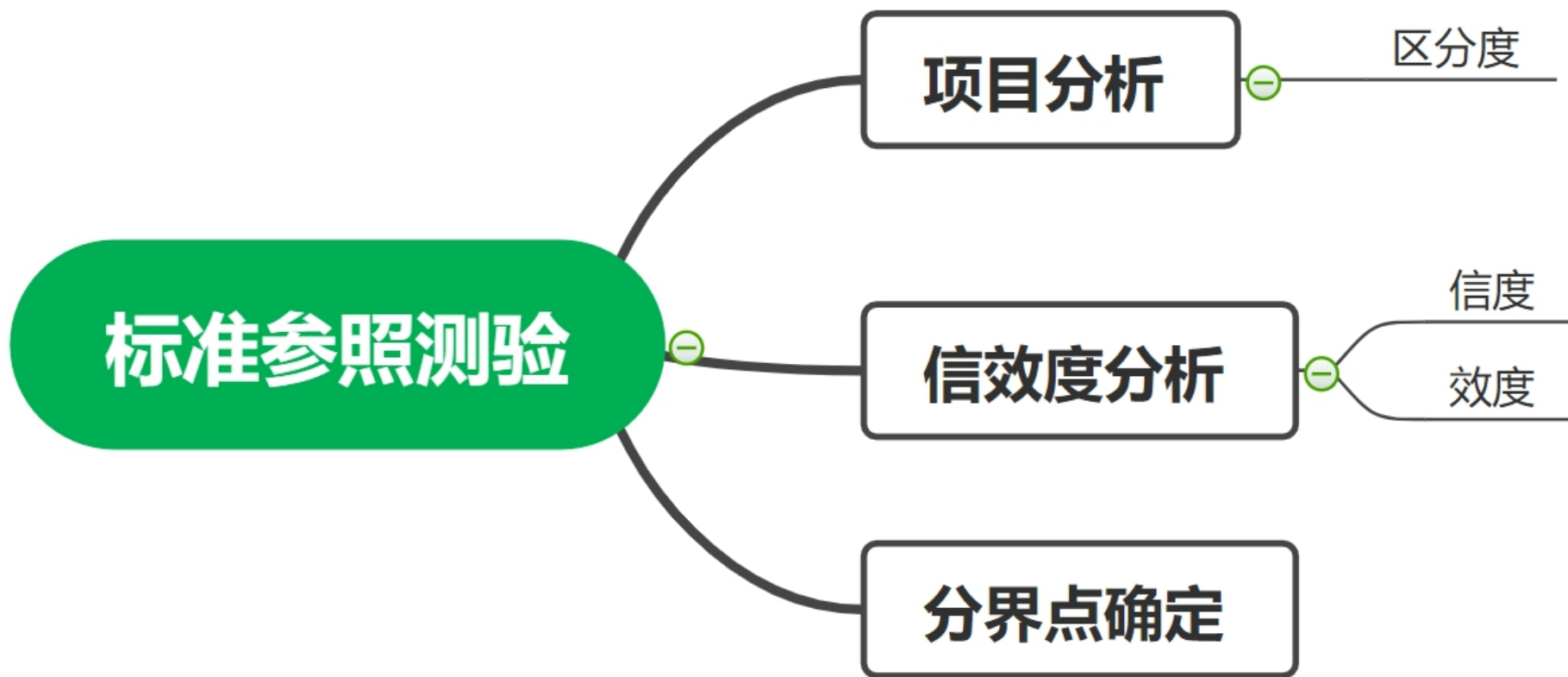


概览

- ◆ 目标参照测验概述（选择）
- ◆ 目标参照测验的项目分析（选择）
- ◆ 目标参照测验的信度与效度（选择）
- ◆ 测验分数的解释——分数分界点的确定（选择）



本章结构





一、目标参照测验概述

- 目标参照测验：又称标准参照测验，是根据某一明确界定的内容范围而缜密编制的测验，并且被试在测验上所得结果，也是根据某一明确界定的行为标准直接进行解释的。
- 通过标准参照测验，可以了解个体在所规定的测量内容上的行为水平，其出发点是个体本身的绝对水平。
- 依据的绝对的标准一般称为“分界点”。



二、目标参照测验的项目分析

(一) 内容范围的确定

- 明确测验内容范围的边界。
- 要明确每一内容范围可分为几类，每一类中又可分为多少更细更小的类别。
- 特定测验目的的确定常为内容范围的界定提供依据。



二、目标参照测验的项目分析

(二) 测验项目的难度与区分度的分析

1. 测验的预测

- 测验编制完成后，须选取一定数量的被试进行预测，由此获得预测数据，然后才能在此数据基础上对项目的难度和区分度进行量化分析。
 - ✓ 前测—后测方法
 - ✓ 已接受教学组—未接受教学组方法
 - ✓ 对照组方法



二、目标参照测验的项目分析

2. 测验项目的难度分析

- 计算：与常模参照测验相同，以通过率表示。
- 有些学者认为，目标参照测验的项目难度分析并不重要，甚至有时并不必要。



二、目标参照测验的项目分析

3. 测验项目的区分度分析

(1) 难度差值

① 掌握组—未掌握组鉴别指数 (D)

- ✓ 分别计算这两组在某项目上的平均通过率，记为 P_A 和 P_B ，
则该项目的鉴别指数为： $D = P_A - P_B$

② 个人获得指数 (DIG)

- 采取前测—后测方法，获得在前测中错误回答某项目而在后测中能够正确回答的被试人数比例。
- 值的范围[0, 1]。
- 反映了经过教学活动之后受益的被试比例。

		后测	
		及格	不及格
前测	及格	a	b
	不及格	c	d



二、目标参照测验的项目分析

(2) 相关系数

- 计算项目得分与测验总分的相关。
- 其值在-1.00和+1.00之间变化。



三、目标参照测验的信度与效度

(一) 信度及其估计

- 通常以相关系数所表示的信度指标不太适用。

1. 分类一致性信度

- 根据某一分数分界点将被试分为及格和不及格两类。

$$P_o = \frac{(a + d)}{(a + b + c + d)}$$

		测验A	
		及格	不及格
测验B	及格	a	b
	不及格	c	d

- 报告分类一致性信度时，必须同时提供测验的分数分界点，以及测验长度等资料。
- ##### 2. 方差分析法——荷伊特信度
- 此法不受测验目的或被试异质性的影响。



三、目标参照测验的信度与效度

(二) 效度及其估计

1. 内容效度

- 目标参照测验一般来说有相对比较确定的内容范围，可用命题细目表表示。
- 评估任一测验的内容效度，都依赖于两个条件：
 - ① 测验有明确界定的内容范围；
 - ② 对测验每一题目的内容效度的分析。
- 方法：专家评定法



三、目标参照测验的信度与效度

2. 效标关联效度

		预测源测验	
		及格	不及格
校标测验	及格	a	b
	不及格	c	d

- 决策效度 = $(a+d)/(a+b+c+d)$



四、分数分界点的确定

(一) 专家判定法

1. Nedelsky 法

- 针对由单选题组成的测验。
- 由专家判断处于临界水平的被试在每一题上有能力排除的错误选择项，计算其正确回答的可能性。
- 求出每一题上正确回答的可能性之和，即为测验分数分界点。

2. Angoff 法

- 由专家直接判断处于临界水平的被试在某测验的每一题目正确作答的可能性（记为 P_i ）。
- 设每一题的满分为 F_i ，计算 $\lambda = \sum F_i P_i$



四、分数分界点的确定

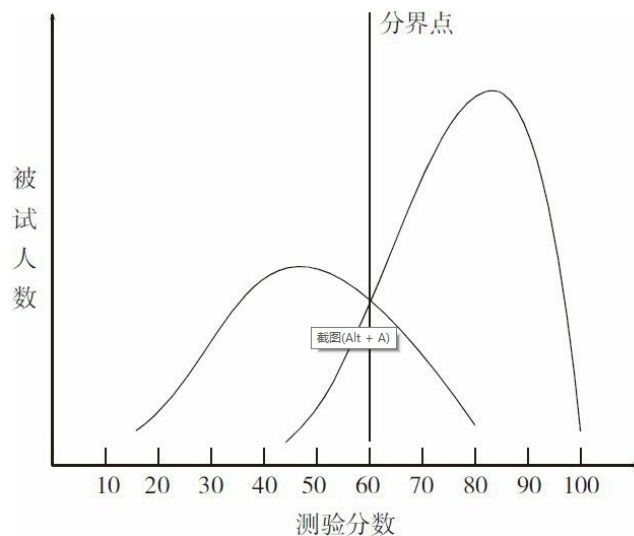
(二) 效标组预测法

1. 临界组法

- 由专家判定一组正处于临界水平的被试。
- 将测验施测于该组被试。
- 计算他们在测验上的平均成绩，可视为分界点的估计值。

2. 对照组法

- 专家选择被试，一组被判定为掌握组，另一组则被判定为非掌握组。不太容易被判定为“掌握”或“非掌握”的被试一概剔除。





真题演练

标准参照测验与常模参照测验最主要的区别是（ ）。(北京师范大学, 专硕, 2015)

- A. 测验的难度要求不同
- B. 适用的信度指标不同
- C. 项目分析指标对样本的依赖程度不同
- D. 分数解释的参照体系不同



真题演练

标准参照测验与常模参照测验最主要的区别是（ ）。(北京师范大学, 专硕, 2015)

- A. 测验的难度要求不同
- B. 适用的信度指标不同
- C. 项目分析指标对样本的依赖程度不同
- D. 分数解释的参照体系不同

答案：D



第十章 测量理论的新发现

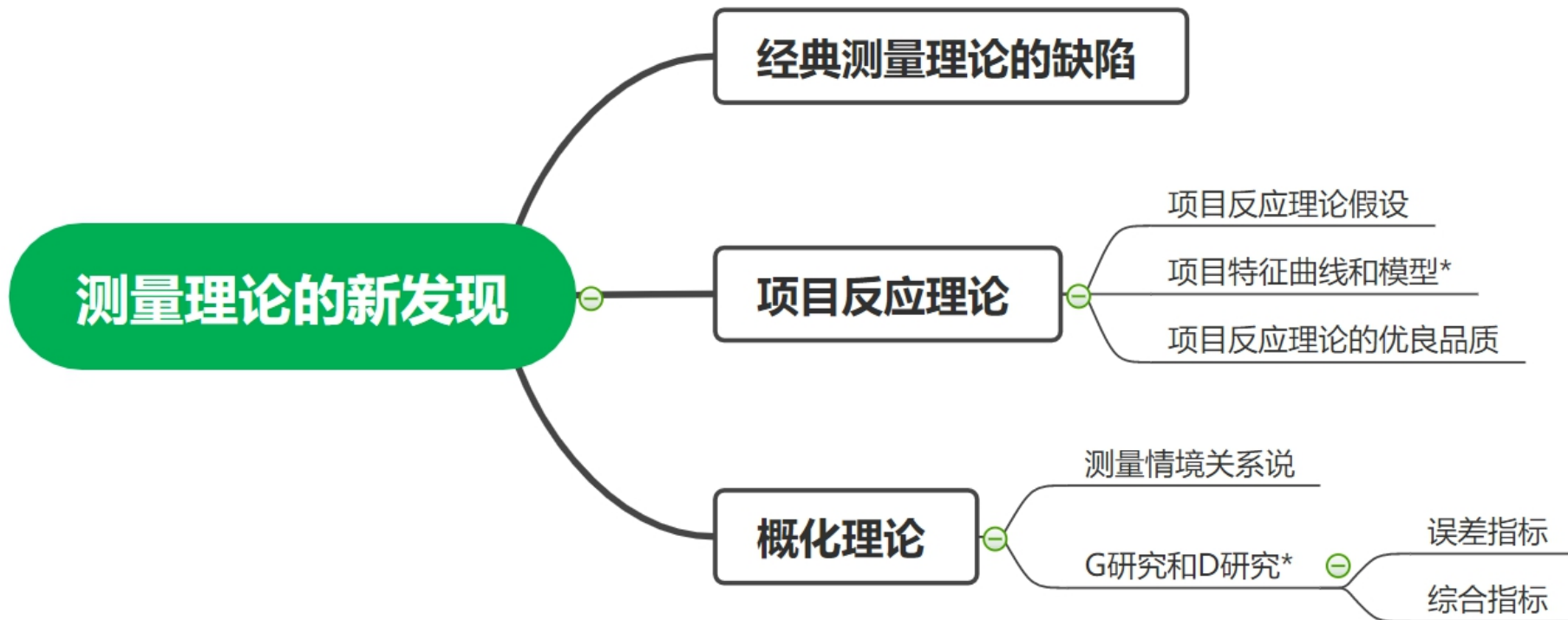


概览

- ◆ 项目反应理论（选择、名词解释）
- ◆ 概化理论（选择、名词解释）



本章结构





一、项目反应理论

经典测量理论的缺陷

1. 信度估计精度不高。
 - 在定义中除原始分数方差可得之外，真分数方差与误差分数方差都是无从求取的；
 - 严格的平行测验是不存在的。
2. 误差指标笼统单一、不精细。
 - 相同的测量对于不同的被试也会有不同的测量误差。
3. 各种参数的估计对样本的依赖性太大。
 - 获得一个对总体有充分代表性的样本是非常困难的。
4. 参数指标之间的配套性较差。
 - 难度和能力不在一个度量参照系上。



一、项目反应理论

(一) 概述

1. 潜在特质

- 制约人的行为的心理特征称为心理特质。由于没有明确的物理或生理形态，又被称为潜在特质。

2. 潜在特质空间

- 对于某一特殊行为的发展起作用的所有潜在特质的集合，称为潜在特质空间。



一、项目反应理论

3. 项目反应理论

- 建立在潜在特质论的基础上，主要研究被试在测验项目上的反应行为与所测潜在特质之间的关系。
- 项目反应理论认为，随着潜在特质 θ 的提高，正确反应该项目的概率 $P(\theta)$ 也会提高。



一、项目反应理论

4. 项目反应理论的假设

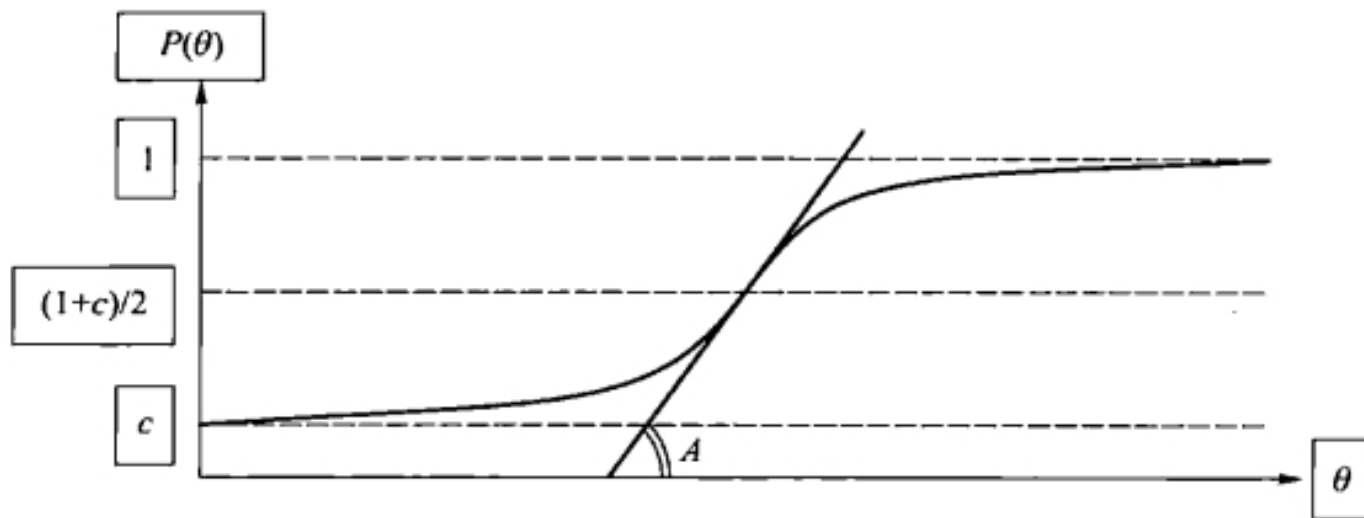
- 大部分IRT是建立在强假设基础上的，主要包括：
 - ✓ **特质空间的单维性假设**：测量结果只取决于一种能力；
 - ✓ **局部独立性假设**：已知能力和项目参数条件下，被试答对某一项目的概率独立于其他项目；
 - ✓ **项目特征曲线假设**：正确反应概率与其能力之间的关系可以用**单调上升函数**表示。



一、项目反应理论

(二) 项目特征曲线

- 指用能稳定反映被试水平的潜在特质变量替代卷面总分作为回归曲线的自变量的曲线，记作 $P(\theta)$ 。
- 用来拟合项目特征曲线的函数，称为项目特征函数(ICF)。

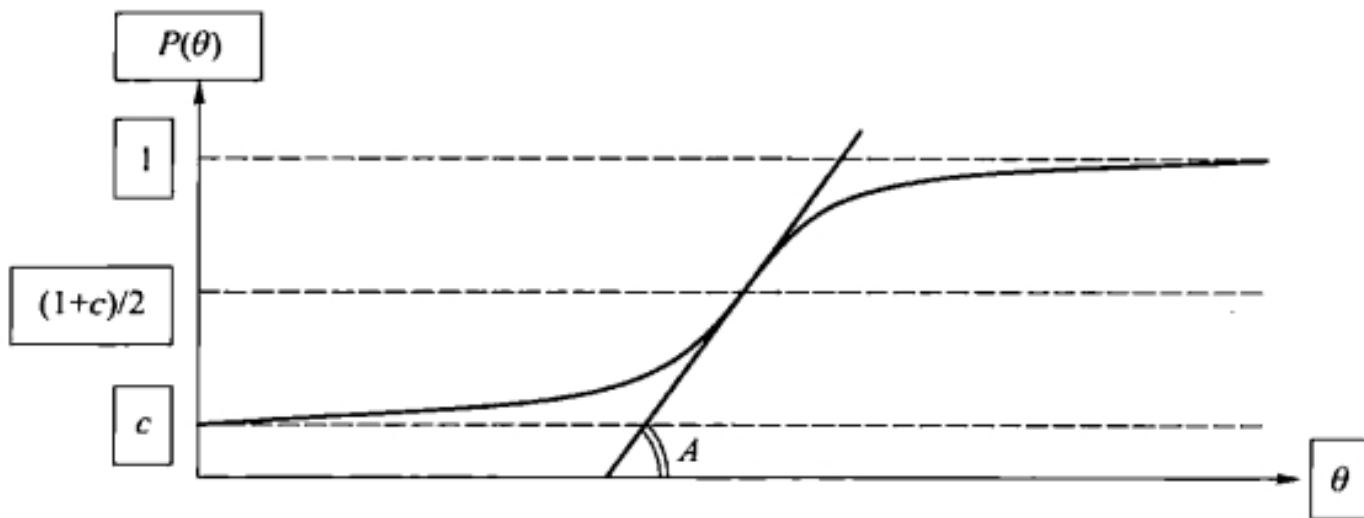




一、项目反应理论

1. 项目特征曲线的形态特点

- Θ 取值范围： $[-\infty, +\infty]$;
- 正确作答的概率，记为 $P(\theta)$ ，取值都在 $[0, 1]$ 内;
- 试题质量较好，则正确作答概率应随被试特质水平的提高而提高。

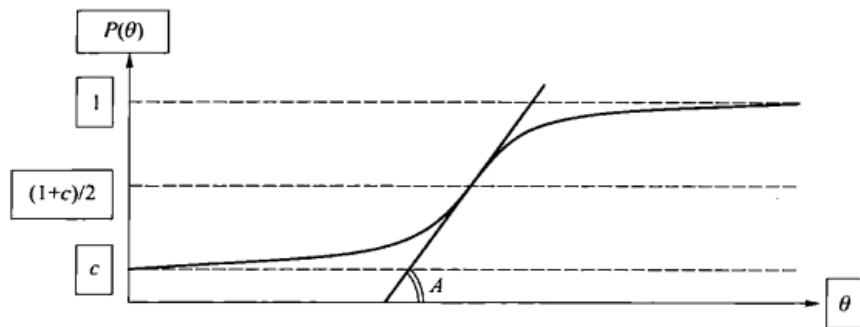




一、项目反应理论

2. 参数

- ① c : 伪机遇水平, 指实际测验中被试纯凭机遇作答而成功的概率。
 $P(\theta)=c$ 是项目特征曲线的下渐近线。
- ② b : 题目难度, 是曲线拐点的横坐标的值。横坐标 $\theta=b$, 纵坐标 $P(\theta)=(1+c)/2$ 的点是项目特征曲线的拐点, 曲线递增的速率在此点由快转慢。
- ③ a : 题目区分度, 刻画题目对被试水平区分能力的高低。 a 值是曲线拐点处切线斜率的函数值。





一、项目反应理论

3. Logistic模型

$$p(\theta) = c + \frac{1 - c}{1 + e^{-1.7a(\theta - b)}}$$

- ① 单参数模型：难度b (a=1 c=0)
- ② 双参数模型：难度b 区分度a (c=0)
- ③ 三参数模型：难度b 区分度a 伪机遇水平c



一、项目反应理论

(三) 项目反应理论的优良性质

1. 参数的跨群体不变性
2. 潜在特质量表的可选择性
3. 参数设计的科学性
 - ① 题目难度参数 b 与被试潜在特质水平参数 θ 定义在同一度量系统上。
 - ② 区分度参数与难度参数相互独立。
 - ③ 伪机遇参数的实证性。
4. 信息函数概念的引进与信息函数的可加性



一、项目反应理论

- 测验项目的信息函数为：
 - $I_i(\theta) = p'_i(\theta)^2 / p_i(\theta) \cdot Q_i(\theta)$
 - 其中， $p'_i(\theta)$ 是 $p_i(\theta)$ 对 θ 的一阶导函数
- 对于一个潜在特质水平值为 θ 的被试，试题 i 施测于他时，所得 θ 值的测量标准误差为：

$$SE_i = [I_i(\theta)]^{-\frac{1}{2}}$$

- 信息函数与试题的区分度成正比，与伪机遇水平成反比，与 θ 减 b 的差的绝对值成反比。
- 测验题目信息函数具有可加性，累加值称为测验信息函数，记为 $I(\theta)$ ： $I(\theta) = \sum_{i=1}^n I_i(\theta)$
- 整个测验的测量标准误差为： $SE(\theta) = [I(\theta)]^{-\frac{1}{2}}$



一、项目反应理论

(四) 项目反应理论的应用

- 计算机化自适应测验的编制
- 认知诊断测量模型发展基础



二、概化理论

(一) 概述

1. 一些测验研究者从深入分析测验误差的来源和结构出发，应用方差分量分析辅助测验研究，创建了从宏观上研究测验性质的新理论——概化理论。
2. **测验情境关系说**：在不同的测验情境关系下，测量误差的结构不同，误差量也不同。
 - ✓ 研究测量必须先研究测验情境关系。
 - ✓ 测验情境关系是由一个测量目标和若干个测量侧面构成的。



二、概化理论

3. 相关概念

- 测量目标
 - ✓ 测量者希望通过测量用测量数据描绘的那些实体。
- 测量侧面
 - ✓ 误差的来源都称为测量侧面。
 - ✓ 一个测量侧面就是某一个方面的测量条件。
 - ✓ 一个测量侧面可以有不同的水平。
- 随机侧面
 - ✓ 指在测量分析中，该侧面内的水平是该侧面所有水平的一个随机样本。该侧面的水平是不固定的。



二、概化理论

- 固定侧面
 - ✓ 在分析中所取水平不是随机样本，而是固定的一个水平，即“**标准化**”。
 - ✓ 固定侧面越多，信度和效度也就越高，测量结果解释、推广的范围就会缩小。
- 多重真分数与多种信度
 - ✓ 测量目标的真分数依赖于测量条件，**在不同测量条件下会有不同的真分数，信度也不相同。**



二、概化理论

4. 对概化理论的理解

- 任何测量都是依赖于特定的测验情境关系的；
- 测验情境关系中的测量目标、测量侧面、测量侧面的水平都是会变化的；
- 上述的变化会引起测验误差的来源、测验误差的大小、真分数的种类以及测验信度的变化，同时测验分数的解释范围也会发生变化。



二、概化理论

(二) G研究与D研究

1. G研究

- 研究者设计的测验情境关系及用一定方法采集的测验数据被称为测验的观察领域。在这观察领域数据上进行的统计分析称为G研究。
- 目的：定量估计观察领域中测量目标的方差以及各个测量侧面所产生的测量误差方差。



二、概化理论

- 步骤

- ① 分解观察数据总方差

- 测量目标主效应方差
- 测量侧面主效应方差
- 各种交互效应方差
 - ✓ 由各测量侧面与测量目标形成的各级交互效应方差
 - ✓ 各测量侧面自己形成的各级交互效应方差

- ② 利用样本方差估计各种效应的期望均方

- 测量目标效应期望均方是测量目标个体差异的描写量
- 各测量侧面效应期望均方是误差描写量。
- 各交互效应期望均方也是一种测量误差。



二、概化理论

2. D研究（决策研究）

- 目的：利用G研究的结果数据，在原设计的测验情境关系范围之内，分析比较各种可能的测验方案，测验工作者可以根据分析结果，结合可能的实施条件**优选实际测验方案**。



二、概化理论

- 误差指标与综合指标
 - 两个比较优劣的误差指标：
 - ✓ **相对误差方差**：所有与测量目标有关的交互效应方差之和
 - ✓ **绝对误差方差**：除测量目标效应方差之外的所有方差之和
 - 测验精度的两个综合指标：
 - **概化系数：G系数**，测量目标效应方差与测量目标效应方差加相对误差方差之和的比，是对常模参照测验分数稳定性程度的度量。
 - **依存系数： ψ 系数**，测量目标效应方差与总效应方差之比，是对目标参照测验分数稳定性和一致性两种程度的度量。



真题演练

在项目反应理论双参数模型中，能够直观地描述被试正确作答概率与被试特质水平关系的是（ ）。(统考，2011)

- A. 区分度与难度
- B. 项目信息函数
- C. 项目特征曲线
- D. 测验信息函数

在项目特征曲线中，反映题目的区分度的是（ ）。(中国政法大学，学硕，2016)

- A. 参数 θ
- B. 项目特征曲线的下渐近线
- C. 项目特征曲线的拐点的横坐标
- D. 项目特征曲线拐点处切线的斜率



真题演练

在项目反应理论双参数模型中，能够直观地描述被试正确作答概率与被试特质水平关系的是（ ）。(统考，2011)

- A. 区分度与难度
- B. 项目信息函数
- C. 项目特征曲线
- D. 测验信息函数

答案：C

在项目特征曲线中，反映题目的区分度的是（ ）。(中国政法大学，学硕，2016)

- A. 参数 θ
- B. 项目特征曲线的下渐近线
- C. 项目特征曲线的拐点的横坐标
- D. 项目特征曲线拐点处切线的斜率



真题演练

在项目反应理论双参数模型中，能够直观地描述被试正确作答概率与被试特质水平关系的是（ ）。(统考，2011)

- A. 区分度与难度
- B. 项目信息函数
- C. 项目特征曲线
- D. 测验信息函数

答案：C

在项目特征曲线中，反映题目的区分度的是（ ）。(中国政法大学，学硕，2016)

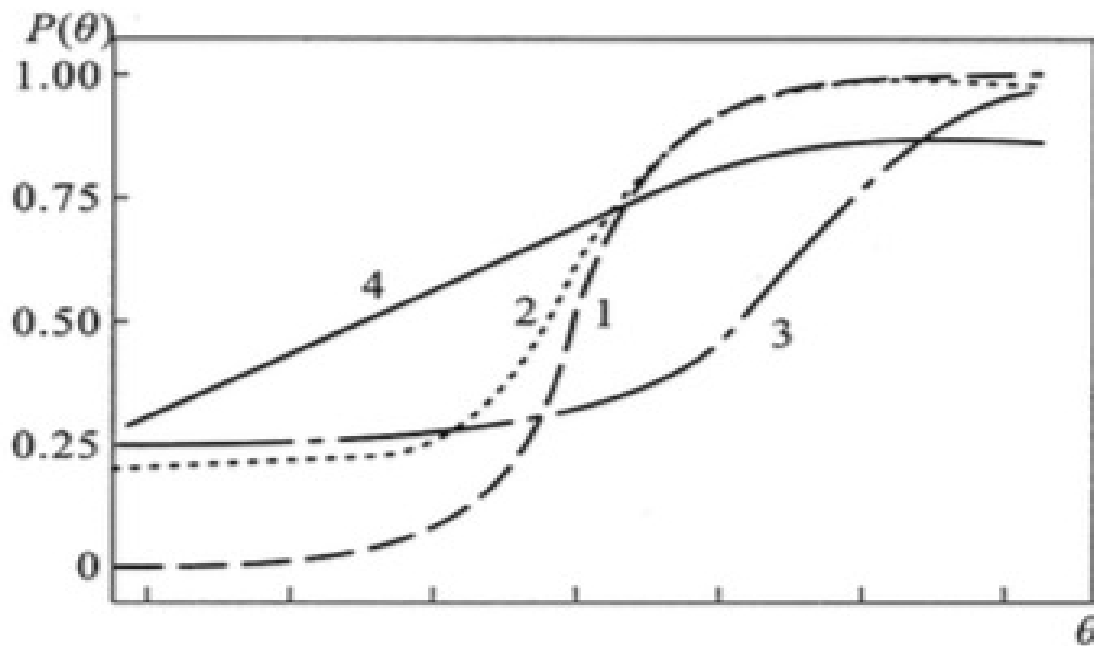
- A. 参数 θ
- B. 项目特征曲线的下渐近线
- C. 项目特征曲线的拐点的横坐标
- D. 项目特征曲线拐点处切线的斜率

答案：D



真题演练

下图是某测验中四个项目的项目特征曲线，其中区分度最小的项目是（ ）。(统考, 2016)



A.项目1

B.项目2

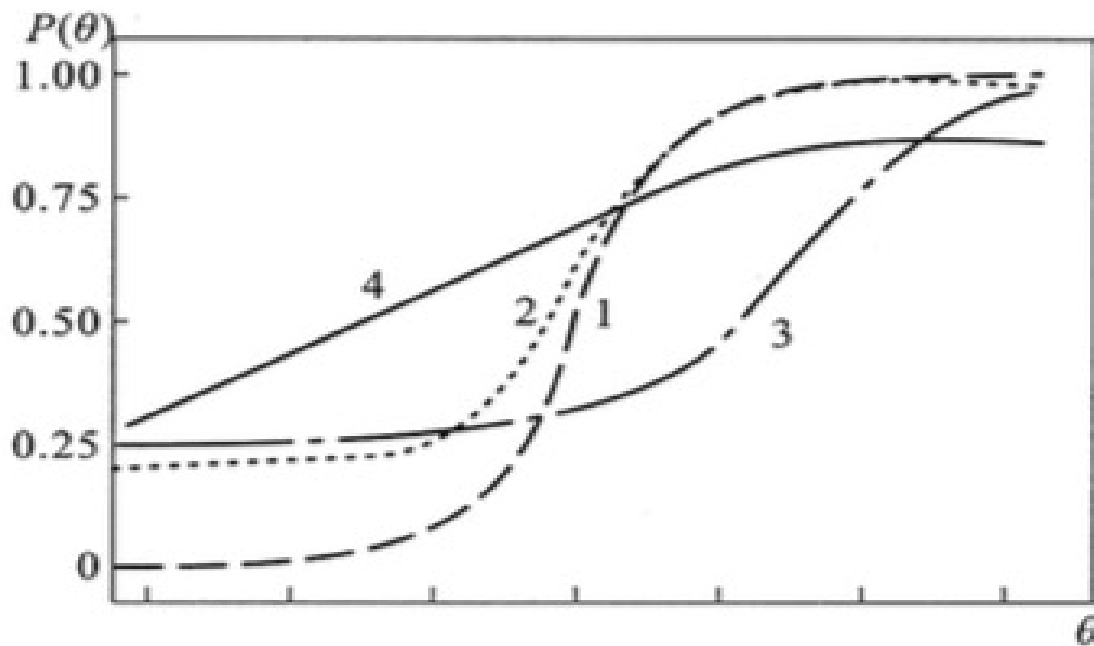
C.项目3

D.项目4



真题演练

下图是某测验中四个项目的项目特征曲线，其中区分度最小的项目是（ ）。(统考, 2016)



A.项目1

B.项目2

C.项目3

D.项目4

答案：D



真题演练

不属于项目反应理论之特性的是（ ）。 （首都师范大学，学硕，2015）

- A. 信息函数的可加性
- B. 潜在特质量表的稳定性
- C. 题目参数的跨群体不变性
- D. 参数设计的科学性

计算机自适应测验设计的主要理论依据是（ ）。 （华中师范大学，学硕，2017）

- A. 概化理论
- B. 经典测验理论
- C. 项目反应理论
- D. 认知诊断



真题演练

不属于项目反应理论之特性的是（ ）。 （首都师范大学，学硕，2015）

- A.信息函数的可加性
- B.潜在特质量表的稳定性
- C.题目参数的跨群体不变性
- D.参数设计的科学性

答案：B

计算机自适应测验设计的主要理论依据是（ ）。 （华中师范大学，学硕，2017）

- A.概化理论
- B.经典测验理论
- C.项目反应理论
- D.认知诊断



真题演练

不属于项目反应理论之特性的是（ ）。 （首都师范大学，学硕，2015）

- A. 信息函数的可加性
- B. 潜在特质量表的稳定性
- C. 题目参数的跨群体不变性
- D. 参数设计的科学性

答案：B

计算机自适应测验设计的主要理论依据是（ ）。 （华中师范大学，学硕，2017）

- A. 概化理论
- B. 经典测验理论
- C. 项目反应理论
- D. 认知诊断

答案：C



真题演练

能够用来考察测量各个侧面变异性的测验理论是（ ）。(统考, 2018)

- A. 概化理论
- B. 经典测验理论
- C. 真分数理论
- D. 项目反应理论

概化理论采用方差分析法分解测量数据的总体方差, 通常把数据总方差分解成（ ）。(多选) (统考, 2013)

- A. 目标主效应方差
- B. 测量侧面效应方差
- C. 各种交互效应方差
- D. 相对误差与绝对误差的方差



真题演练

能够用来考察测量各个侧面变异性的测验理论是（ ）。 （统考，2018）

A.概化理论

B.经典测验理论

C.真分数理论

D.项目反应理论

答案：A

概化理论采用方差分析法分解测量数据的总体方差，通常把数据总方差分解成（ ）。 （多选） （统考，2013）

A.目标主效应方差

B.测量侧面效应方差

C.各种交互效应方差

D.相对误差与绝对误差的方差



真题演练

能够用来考察测量各个侧面变异性的测验理论是（ ）。 （统考，2018）

A.概化理论

B.经典测验理论

C.真分数理论

D.项目反应理论

答案：A

概化理论采用方差分析法分解测量数据的总体方差，通常把数据总方差分解成（ ）。 （多选） （统考，2013）

A.目标主效应方差

B.测量侧面效应方差

C.各种交互效应方差

D.相对误差与绝对误差的方差

答案：ABC



真题演练

在概化理论中，D研究是指（ ）。(统考，2015)

- A. 概化研究
- B. 决策研究
- C. 交叉研究
- D. 嵌套研究

在测量的概化理论中，G系数一般是指（ ）。(首都师范大学，学硕，2014)

- A. 测量目标效应方差与总效应方差之比
- B. 测量目标效应方差与相对误差方差之比
- C. 测量目标效应方差与它加上相对误差方差的和之比
- D. 相对误差方差与总效应方差之比



真题演练

在概化理论中，D研究是指（ ）。(统考，2015)

- A.概化研究
- B.决策研究
- C.交叉研究
- D.嵌套研究

答案：B

在测量的概化理论中，G系数一般是指（ ）。(首都师范大学，学硕，2014)

- A.测量目标效应方差与总效应方差之比
- B.测量目标效应方差与相对误差方差之比
- C.测量目标效应方差与它加上相对误差方差的和之比
- D.相对误差方差与总效应方差之比



真题演练

在概化理论中，D研究是指（ ）。(统考，2015)

- A.概化研究
- B.决策研究
- C.交叉研究
- D.嵌套研究

答案：B

在测量的概化理论中，G系数一般是指（ ）。(首都师范大学，学硕，2014)

- A.测量目标效应方差与总效应方差之比
- B.测量目标效应方差与相对误差方差之比
- C.测量目标效应方差与它加上相对误差方差的和之比
- D.相对误差方差与总效应方差之比

答案：C

THANKS

- 期待下次相遇 -



高途学院APP下载



高途学院公众号

