题目：曝光控制约束条件下的几种CAT终止规则研究

实验内容：

为了充分研究CAT在不同变长终止规则下的表现情况。本文预计采用控制变量法设计两类试验。

第一类是确定选题策略，令终止规则为变量来分析不同终止规则在同一选题策略下的表现情况，通过比较所得结果可得某一选题策略最适配的变长终止规则。

第二类是确定终止规则为定长终止规则，令选题策略和题量为变量来分析不同题量在不同选题策略上的表现情况，再将表现最好的定长题量与不定长终止规则的实验结果进行比较，判断出优劣。

基于此比较所有实验结果，最终得出着重考虑曝光控制情况下的变长终止规则表现情况，比较出优劣。

逻辑流程：

图示

描述已自动生成

选题策略：a分层，b分层，最大信息量

变长终止规则：最小信息量，迭代逼近

题库题目数500 受试者人数50

定长终止规则题量 5 10 20，选题策略同上

1. 最大信息量选题策略的原理是，根据当前受试者的预估能力来计算题库中剩余题目所含的信息量，从而为受试者选择下一个题目，信息量的计算方法是抽样方差求导数。

图示

描述已自动生成

1. 按a分层选题策略的原理是，首先根据题目区分度a将题库中所有题目分为t层，再将整个测试分为t个阶段，在测试的第t段一一选择个与当前预估能力值均衡的题目，让受试者继续作答，其中等于受试者测试长度。随后重复前一步实验，最终得出结果[19]。

图示

描述已自动生成

（3）按b分层选题策略是指，首先根据难度数b将题库中所有题目分为q层，再将整个测试分为q个阶段，在测试的第q段逐一选择个与当前预估能力值均衡的题目，让受试者继续作答，其中要等于受试者测试长度。 随后重复前一步实验，最终得出结果。

B分层流程类似a分层

1. 最小信息量终止规则基本原理是，设定一个题库中剩余题目的最小信息量阈值（根据每道题目的信息量量级设置），当受试者做了一定题目得出能力值后计算剩余题目所含信息量，当该信息量小于等于所设最小信息量阈值，则停止试验。
2. 能力值与估计误差终止规则指，设定一个终止误差，给出一个能力值估计，使用牛顿迭代法，受试者每做一定题目就迭代计算新的能力值，比较预估与计算的误差，当两者误差小于等于所设终止误差，则终止试验。

受试者初始能力估计流程：

图示

描述已自动生成

（两参数的题库构建，c=0）

参考论文p50~p56

（受试者答对某题的概率）

其中，是考生在测验所测的特质上的水平；（代码敲好了）

D=1.702是用Logistic函数趋近正态肩形曲线时所使用的常数因子；

分别是代表试题区分度、难度和猜测度的项目参数[18]。

题目信息量计算公式，最大信息量最小信息量用

其中，代表试题在能力为上所提供的信息，

为在点上的值导数，

为能力为在试题上的试题反应函数，



评判终止规则优劣指标：

3.2构建方法

本文的评价模型内具体指标有，能力估计准确性（）（普通方差），模型策略稳定性（）（减初始值的方差） ，人均用题数 （），题目曝光度（），卡方统计量（），未使用试题量 ，其中判断受试者能力值是否与实际值相符，判断同一受试者不同试验中测出的能力值是否稳定，判断受试者使用题目的多少来评判策略的效率，题目曝光度，未使用试题量和用于判断题库的曝光程度，即题库是否存在泄漏题目的风险。

公式：p57~p58

题库构建时a选择对数正态分布，b选择均匀分布