假设检验(知识点为单位)

认知诊断

输入 1：学生做题矩阵 $R_{students \times items}$，$R(i,j) = 0/1$ 表示第 i 个学生做第 j 道题目的对错

输入 2：$Q_{items \times skills}$ 矩阵，$Q(i,j) = 0/1$，表示第 i 个道题目是否考察第 j 个知识点

认知诊断模型有很多，如简化的非补偿性认知诊断模型 $DINA$、或者 $DINO$ 、$GDINA$ 等

以 $DINA$ 为例，项目反应函数

$$P(X_{ij}=1|\alpha_i)=(1-s_j)^{\eta_{ij}} g_j^{1-\eta_{ij}}, \quad \eta_{ij}=\prod_{k=1}^{K} \alpha_{ik}^{q_{jk}}.$$

输出 1：预测学生对 $skills$ 个知识点是否掌握

输出 2：估计 $DINA$ 类模型中的猜测参数 $guess$，和失误参数 $slip$

假设检验

符号定义

- 学生数量：I
- 题目数量：J
- 知识点数量：K
- 作答矩阵：$R_{I \times J}$，其中 r_{ij} 表示第 i 个学生对第 j 道题目的作答对错，正确为 1，错误为 0
- Q 矩阵：$Q_{J \times K} = (q_{jk})_{J \times K} = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_J \end{pmatrix}$，学生掌握情况矩阵

  $\beta_{I \times K} = (\beta_{ik})_{I \times K} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_I \end{pmatrix}$
- $q_j = (q_{j1}, q_{j2}, \ldots, q_{jK})$ 为第 j 题对 K 个知识点的考察情况，考察记为 1，未考察记为 0

- $\beta_i=(\beta_{j1},\beta_{j2},\ldots,\beta_{jK})$ 为第 i 个学生对 K 个知识点的掌握情况，掌握记为 1，未掌握记为 0
- $\beta_i \succeq q_j$ 表示 β_i 中的每个分量不小于 q_j 中的对应分量，即对 $\forall k$ 都有 $\beta_{ik}\geq q_{jk}$
- $Q_{j;1,\ldots,K}$ 表示 Q 矩阵的子矩阵，特别的表示 Q 矩阵的第 j 行
- 定义运算 $Q\setminus Q_{[j;1,\ldots,K]}=\begin{pmatrix} q_1 \\ \ldots \\ q_{i-1} \\ q_{i+1} \\ \vdots \\ q_J \end{pmatrix}_{(J-1)\times 1}$ ，表示 Q 矩阵去除第 j 行 q_j 向量，记 $Q_{\setminus[j,:]}=Q\setminus Q_{[j;1,\ldots,K]}$
- 同理 $R\setminus R_{[1,\ldots,I;j]}=(r_1,\ldots,r_{j-1},r_{j+1},\ldots,r_J),$，表示作答矩阵 R 去除第 j 列作答向量，记 $R_{\setminus[:,j]}=R\setminus R_{[1,\ldots,I;j]}$
  - 借用集合中的定义： $A\setminus B$ 表示 $\{x:x\in A \quad and \quad x\notin B\}$，推到矩阵上的定义

思路

以第 j 道题目为例

1. 第一步： $DINA$ 模型输入 Q,R ，输出第 j 题的猜测参数 g_j，失误参数 s_j，每个学生的掌握情况 β
2. 第二步：设计方法推出某个 q_{jk} 可能存在问题，并且能推出是属性缺失问题，还是属性冗余问题 (也可以直接遍历所有题目，不用判断是否缺失冗余)
3. 第三步：$DINA$ 模型输入 $Q_{\setminus[j,:]},R_{\setminus[:,j]}$ ，输出 $J-1$ 道题目的猜测参数 $\tilde{g}$ 和失误参数 $\tilde{s}$，以及每个学生的掌握情况 $\tilde{\beta}$。此步骤作用是为了降低第 j 题 q_j 向量错误导致其他参数估计的误差。
4. 第四步：确定假设检验问题
   1. 情形一：判断出来是属性缺失问题
   2. 情形二：若判断出来是属性冗余问题
   3. 情形三：同时存在属性缺失和冗余问题，首先按.. 待推导
5. 第五步：选择样本
   1. 若为情形一缺失情况
      1. 则建立假设：

$$H_0: q_{jk}=0, \leftrightarrow H_1: q_{jk}=1$$

2. 选择满足以下规则的样本，输入某道题目的 q_j 向量，指定第 k 个知识点，以及 $\tilde{\beta}$

$$T = \{i : \tilde{\beta}_i \succeq q_j, \quad \text{且} \tilde{\beta}_{ik} = 0\}$$

3. 建立统计量：根据 T 中的学生，计算作答第 j 题的错误数量 X,

$$X = n_T - \sum_{i \in T} r_{ij}$$

4. 原假设成立时，样本 T 中的学生具有做对 q_j 考察模式题目的掌握模式，因此做第 j 题错误只能是失误，因此错误概率为 $DINA$ 估计的失误参数。但 $DINA$ 模型的输入为 $Q_{\setminus[j,:]}, R_{\setminus[:,j]}$，输出无第 j 题的失误参数，因此采用其他题目的平均作为估计，$\tilde{s}_j = \sum_{l=1}^{J-1} \tilde{s}_l$，统计量服从二项分布：

$$X \sim B(n_T, \tilde{s}_j), \quad P(X = x) = C_{n_T}^{x} \tilde{s}_j^x (1 - \tilde{s}_j)^{(n_T - x)}$$

5. 原假设不成立时，$q_{jk} = 1$，即第 j 题考察了第 k 个知识点，而 T 中的学生均未掌握第 k 个知识点，因此做错的数量会更多，X 有偏大的趋势，拒绝域形式为 $[c, +\infty]$。即

$$W_1 = \{(r_{1j}, r_{2j}, \ldots, r_{n_T j}) : X > c\} = \{n_T - \sum_{i \in T} r_{ij} > c\}$$

6. 给定置信度 α 进行检验：

$$P(X \geq c) = \sum_{x=c}^{n_T} P(X = x) \leq \alpha$$

7. 【缺失属性例子 1】
    1. 真实 $q_{j_{\text{真}}}$ 向量为 $[1, 1, 0, 0]$,
    2. 当下错误 q_j 向量为 $[1, 0, 0, 0]$
    3. 判断出第 j 题是缺失情况，则 $H_0 : q_{j2} = 0, \leftrightarrow H_1 : q_{j2} = 1$
    4. 根据 $Q_{\setminus[j,:]}, R_{\setminus[:,j]}$ 估计了除第 j 题外的 $\tilde{g}, \tilde{s}$，因此第 j 题的估计 $\tilde{s}_j = \sum_{l=1}^{J-1} \tilde{g}_l$，估计了学生的掌握模式 $\tilde{\beta}$
    5. 若第 $k = 2$ 个知识点存在缺失属性的情况
    6. 根据错误的 q_j 筛选样本 $T = \{i : \tilde{\beta}_i \succeq q_j, \quad \text{且} \tilde{\beta}_{i2} = 0\}$
        1. 筛选的掌握模式属于以下类型：
            1. $[1, 0, 0, 0]$
            2. $[1, 0, 1, 0]$
            3. $[1, 0, 0, 1]$

      4. $[1,0,1,1]$
   7. 计算统计量：样本 T 做第 j 题的错误数量 X，如果假设正确，则做错只可能是失误，$X \sim B(n_T, \tilde{s}_j)$，例如 $n_T = 30, \tilde{s}_j = 0.2, X = 10, \alpha = 0.05$
   8. 则计算累计概率 $P(X \geq 10) = 0.0256 < 0.05$
   9. 因为错误题目太多了，拒绝原假设，则修改 q_j 矩阵为 $[1,1,0,0]$
  8. 【缺失和冗余同时存在的例子】
2. 若为情形二冗余情况
  1. 则建立假设：

$$H_0 : q_{jk} = 1, \leftrightarrow H_1 : q_{jk} = 0$$

  2. 选择满足以下规则的样本

$$S = \{i : \tilde{\beta}_i \succeq q_j, \quad 且 \tilde{\beta}_{ik} = 0\}$$

  3. 建立统计量：根据 S 中的学生，计算作答第 j 题的正确数量 Y，

$$Y = \sum_{i \in S} r_{ij}$$

  4. 原假设成立时，样本 S 中的学生掌握模式 β_i 相比 q_j 少了第 k 个知识点，所以做第 j 题时一定答错，如果对了那么只可能是猜对！因此猜测概率为 $DINA$ 估计的猜测参数。但 $DINA$ 模型的输入为 $Q_{\setminus[j,:]}, R_{\setminus[:,j]}$，输出无第 j 题的猜测参数，因此采用其他题目的平均作为估计，$\tilde{g}_j = \sum_{l=1}^{J-1} \tilde{g}_l$，统计量服从二项分布：

$$X \sim B(n_S, \tilde{g}_j), \quad P(X = x) = C_{n_S}^x \tilde{g}_j^x (1 - \tilde{g}_j)^{(n_S - x)}$$

  5. 原假设不成立时，$q_{jk} = 0$，即第 j 题未考察了第 k 个知识点，而 S 中的学生此时做第 j 题一定做对的，Y 有偏大的趋势，拒绝域形式为 $[c, +\infty]$。即

$$W_2 = \{(r_{1j}, r_{2j}, \ldots, r_{n_S j}) : Y > c\} = \{n_T - \sum_{i \in S} r_{ij} > c\}$$

  6. 给定置信度 α 进行检验（做对的人太多说明这题没考这么多知识点）

$$P(Y \geq c) = \sum_{x=c}^{n_S} P(X = x) \leq \alpha$$

思考

问题 1

如何判断第 j 道题目是缺失还是冗余?