**Cognitive Diagnosis Models**

[Robert A. Henson](http://www.uncg.edu/~rahenson/myvita.htm)

**1 Introduction**

If the goal of a test is to accurately measure an examinee’s general ability, indices exist that can provide an indication of the value of each item. However, modern methods for skills diagnosis are interested in determining mastery on K dichotomous skills rather than assessing general ability, and therefore typical indices of a “good” item cannot directly apply. A new set of indices used to indicate a good item must be developed. First, a brief description of cognitive diagnosis models is provided, followed by the problems of deﬁning a good item when using skills diagnosis models. Then, we deﬁne two indices that can be used to indicate the value of an item and show their relationship to estimation of the examinees’ mastery on a set of skills using a Monte Carlo simulation study.

As opposed to estimating an examinee’s ability along a continua, skills diagnostic models (cognitive diagnosis models, CDMs) typically estimate a student’s mastery proﬁle and therefore provide those skills that an examinee has mastered or has not mastered. In doing this, the probability of a correct response is modeled as a function of mastery of the set of K skills (the mastery proﬁle, α). The general basis underlying such models is that in order to correctly respond to an item, one must have mastered a basic set of skills, if any of these skills have not been mastered, the chances of a correct response is lowered.

Because examinees are characterized by a set of dichotomous skills, most cognitive diagnosis models can be directly compared to latent class models. Speciﬁcally, most CDMs are constrained latent class models where each class is deﬁned by the mastery proﬁle (these have also been called multiple latent class models). All examinees with the same mastery proﬁle (i.e., examinees in the same latent class) are assumed to have the same expected response pattern. Macready and Dayton (1977) were among some of the ﬁrst to discuss such mastery latent class models using only one dichotomous trait to measure mastery of a test domain. In addition, Rindskopf (1983) suggests using latent class analysis with particular constraints placed on the item probabilities. Later, Haertel (1989) parameterized a simple model called the Binary Skills Model, which was later called the DINA model (Deterministic Inputs, Noisy ”And” gate model) by Junker and Sijtsma (2001). The following sections describe some common models in more detail.

**1.1 Current Cognitive Diagnosis Models**

As was brieﬂy mentioned, CDMs deﬁne the probability of a correct response based on an attribute mastery proﬁle. Because all individuals with the same attribute mastery proﬁle are expected to perform in the same way on any given item, CDMs are similar to latent class models. Speciﬁcally, CDMs are a special case of constrained latent class models where classes are deﬁned by the attribute mastery proﬁles (see references here). All individuals that have mastered the same set of required attributes for any given item will have the same probability of a correct response for that item (i.e., probabilities are constrained to be same across all attributes that have the same proﬁle for an items required items). For all models discussed within this paper, a Q-matrix is an item by attribute indicator matrix that deﬁnes which attributes that must be mastered to have a high probability of a correct response.

A brief example is provided to better describe the Q-matrix using three items that may be included in a test used to measure basic math ability. In Table 1 we provide the three possible items.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | Table 1: | Three Example Items for a Test Measuring Basic Math |  |  |  |  | | --- | --- | --- | |  |  |  | | 1. | 4+3-1= | ? | | 2. | 4/2= | ? | | 3. | 4\*2-5+2= | ? | |  |  |  | |  |  |  | |

IRT or CTT would generally be used to estimate an examinee’s basic math ability for each examinee and therefore as math ability increases for that examinee the probability of a correct response would also increase. However, in this case, using CDMs one would assume that only mastery of a few basic skills is required to correctly respond to any item. In answering these three questions basic math could be reduced to the mastery of addition, subtraction, multiplication, and division. CDMs assume that an individual would need to have mastered both addition and subtraction to have a high probability of a correct response for Item 1, whereas Item 2 would need only division and Item 3 would need mastery of addition, subtraction, and multiplication. The Q-matrix summarizes this information in an Item by Attribute indicator matrix. Table 2 provides the Q-matrix for this example, where a one means that the attribute is required and a zero means that the attribute is not required.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | Table 2: | Example Q-matrix Entries |  |  |  |  |  |  | | --- | --- | --- | --- | --- | |  |  |  |  |  | | Item | Add | Sub | Mul | Div | |  |  |  |  |  | | 1. | 1 | 1 | 0 | 0 | | 2. | 0 | 0 | 0 | 1 | | 3. | 1 | 1 | 0 | 1 | |  |  |  |  |  | |  |  |  |  |  | |

Given the Q-matrix and an examinee’s attribute pattern, CDMs deﬁne the probability of a correct response. However, how each model deﬁnes the probability of a correct response can differ. For example, a simple model may assume that the only way to have a high probability of a correct response is to have mastered all required attributes and otherwise the probability of a correct response is the probability that the correct response is guessed. Other models assume that for each attribute mastered, the probability of a correct response increases.

The functional relationship between the Q-matrix and attributes mastered with the probability of a correct response can be categorized as one of two types of CDMs: noncompensatory and compensatory. The following sections describe models classiﬁed as each.

**1.2 Noncompensatory Models**

Noncompensatory models have been divided into two different types of models, conjunctive and disjunctive models. Conjunctive models are deﬁned such that knowing all required attributes for an item increases the probability of a correct response above and beyond what would be expected from the sum of their individual effects. Math items provide a typical example when this may be true. If a math item required two different skills such as addition and subtraction (e.g., 4+3-1=?), an examinee knowing only one of these two attributes would still have to guess the correct answer. However, by knowing both addition and subtraction this problem could be directly solved. In contrast, disjunctive models allow for only minor increases in the probability of a correct response when additional attributes are mastered. In this case, it is as though there are multiple strategies to solve an item. If an examinee were to master a subset of required skills, which are needed for one strategy, the probability of a correct response would be high, but there would be little to no gain if additional attributes were also mastered. To better describe the differences between conjunctive and disjunctive models, descriptions of some common models are provided.

**1.2.1 Conjunctive Models**

Among some of the simplest conjunctive models is the DINA (Deterministic Input; Noisy ”And” Gate) model. In the DINA model items divide the population into two classes, those who have all required attributes and those who do not, which is indicated by the latent variable, ξij. The value ξij only equals one for those examinees who have mastered all required attributes and is zero otherwise. Speciﬁcally, for examinee i for item j.

     K∏   qjk
ξij =    αik ,
    k=1

where αj is a (K x 1) 0/1 vector such that the kth element for the ith examinee, αik, indicates mastery, or non-mastery, of the kth attribute. Once the value of ξij is known the probability of a correct response for examinee j for correctly responding to item j is deﬁned by sj and gj. Where sj is the probability of incorrectly answering an item when in fact all required attributes have been mastered (or “slipping” up) and gj is the probability of correctly responding to an item when all required attributes for that item have not been mastered, a “guessing” parameter.

|  |  |
| --- | --- |
| sj = P(Xij = 0∣ξij = 1) | (1) |

|  |  |
| --- | --- |
| gj = P(Xij = 1∣ξij = 0) | (2) |

Given the jth item’s parameters and ξij, the probability of a correct response can be written as

|  |  |
| --- | --- |
| ξij(1-ξij) P (Xij = 1∣ξij,sj,gj) = (1- sj) gj    . | (3) |

In this model one additional constraint is the (1 - sj) > gj and therefore an examinee mastering all required attributes, ξij = 1, has a higher probability than an examinee who has not mastered all required attributes, ξij = 0. Notice that by deﬁning the DINA model in this way, the probability of a correct response is only high when an examinee has mastered all required attributes. If only a subset of the required attributes have been mastered there is no increase. Junker and Sijstma (2001), Tatsuoka (2002) and de la Torre and Douglas (2003) discuss methods for estimation of the DINA model, which implement MCMC algorithms. The DINA model has also been considered by Macready and Dayton (1977), Haertel (1989), and Doignon and Falmagne (1995). As the latent class part of a mixture model, in which other examinees follow a latent trait model, Yamamoto (1987) employs a similar formulation.

The DINA is the most extreme example of a conjunctive model, because there is a large increase in the probability of a correct response when all required attributes have been mastered but there is no increase (or effect) in the probability of a correct response when any subset of the required attributes have been mastered. The DINA has been considered too restrictive in that all examinees who lack at least one required attribute are assumed to have the same probability of a correct response. For certain situations it may seem more realistic that the probability of a correct response increases for each required attribute that has been mastered. One model that addresses this concern is the Reparameterized Uniﬁed Model (RUM; Hartz et al 2005).

The RUM deﬁnes the probability of a correct response for item j as a function of an examinee’s mastery proﬁle αi and a continuous ability ηi. The examinee parameter η indicates an examinee’s ability on all other skills that are required to correctly answer the item, but have not been speciﬁed in the q-matrix. Given both examinee parameters the RUM deﬁnes the probability of a correct response as:

|  |  |
| --- | --- |
| K∏    q (1-α ) P (Xij = 1∣αi,ηi) = π*j   rj*k jk    ik Pcj(ηi),                     k=1 | (4) |

Here, the item parameter πj\* is deﬁned as the probability of a correct response to item j assuming that all required attributes have been mastered and assuming that the examinee is knowledgable in all residual attributes that are required for the item (i.e., examinee has a very high ηi). The rjk\* parameters, are constrained such that 0 ≤ rjk\*≤ 1, and indicate the proportional amount that the probability of a correct response to item j is reduced if the kth required attribute (qjk = 1) has not been mastered. Lastly, cj is equal to the negative of the difficulty parameter of a Rasch model. That is,

|  |  |
| --- | --- |
| 1.701(ηi+cj) Pcj(ηi) =--e-1.701(η-+c).         1+ e     i  j | (5) |

Notice, as an examinee becomes less proﬁcient in the required residual skills captured by ηi, Pcj(ηi) approaches zero and thus proportionally reduces the probability of a correct response (i.e., it acts as though it is another r\*). Although the full model is most appealing because of its ability to indicate the extent to which the q-matrix has fully been speciﬁed based on the value of cj, this parameter has been difficult to estimate. For that reason we focus on a reduced version of the RUM (RedRUM) where Pcj(ηi) = 1. Therefore the RedRUM will be addressed throughout the rest of the model and is deﬁned as:

|  |  |
| --- | --- |
| K∏    q (1-α ) P (Xij = 1∣αi,ηi) = π*j   r*jk jk    ik .                    k=1 | (6) |

Notice again that because the RedRUM is multiplicative and there is an rij\* for each required attribute, the probability of a correct response decreases for each attribute that has not been mastered. In addition, the impact of mastering all attributes is greater than the sum of the effects of having only mastered any one required attribute. It should be noted that another model with a similar goal (although not indentical) to the RedRUM is the NIDA (Noisy Input; Deterministic “And” Gate, Maris, 1999) model, where the effect of not mastering any one attribute is ﬁxed across all items that require that attribute.

**1.2.2 Disjunctive Models**

In general, all conjunctive models allow for an increase in the probability that is largest when all other required attributes for that item have been mastered. As an alternative, disjunctive models assume that mastery of additional attributes provide little to no gain once a subset of the required attributes have been mastered. Templin and Henson (2006) provide an example in psychology when such a model would be reasonable. Speciﬁcally, an item such as, ”I am ashamed of the things I’ve done to obtain money for gambling”, which was used to measure pathological gambling. Such an item could have a positive response for more than than one reason. Although this item can be thought to measure the extent that a person has committed illegal acts, an individual who has relied on others to ﬁnance his or her gambling may also have a high probability of endorsing such an item. In addition, the probability of endorsing this item may only marginally increase for an individual who has done both.

In their paper, Templin and Henson (2006) describe the DINO (Deterministic Input; Noisy “Or” Gate) model. The DINO, much like the DINA, models the probability of a correct response as a function of a slipping parameter, sj, and a guessing parameter, gj. However, instead of deﬁning ξij they use the parameters ωij. The latent variable ωij is deﬁned as

|  |  |
| --- | --- |
| K∏         q ωij = 1 -  (1- αik)jk,          k=1 | (7) |

which is an indicator of whether the ith examinee has mastered at least one of the required attributes for the jth item. Therefore, ωij = 1 for any examinee having mastered one or more of the item required attributes and ωij = 0 for an examinee who has not mastered any of the required attributes. Given ωij the probability of a correct response is deﬁned as:

|  |  |
| --- | --- |
| ωij (1-ωij) P(Xij = 1∣ωij,sj,gj) = (1- sj) gj    . | (8) |

As in the DINA, examinees are divided in to two groups, however once one of the required attributes has been mastered the probability of a correct response increases from gj to (1 - sj). The probability of a correct response does not increase for examinees who have mastered more than one of the required attributes and therefore, this examinee would perform worse than what would be expected if one were to sum the effects of learning each attribute separately.

**1.3 Compensatory Models**

All noncompensatory models are similar in that the effect of mastering several of the required attributes is better (or worse) than one would expect by summing the effects of mastering each attribute separately. As an alternative, compensatory models are such that mastering an attribute increases the probability of a correct response in the same way regardless of what additional attributes have been mastered. Speciﬁcally, the probability of a correct response having mastered all required attributes is related to the sum of the effects of mastering any one attribute.

Probably one of the simplest examples of such a model is the Compensatory RUM. the compensatory RUM is a multidimensional IRT model where instead of deﬁning ability as a set of continuous ability measures, ability is measured as a proﬁle of 0/1 attributes, α. Speciﬁcally, the compensatory RUM is deﬁned as:

|  |  |
| --- | --- |
| ∑                e  Kk=1r*jkqjk-bj P (X  = 1∣α) = ----∑K---r*qjk-b-              1+ e  k=1 jk    j | (9) |

Notice that in the compensatory RUM, the effect of mastering each attribute is deﬁned by rjk\*. For each additional attribute that is mastered, an additional rjk\* is added, however there is no added impact by knowing several required attributes for that item