Etude de Classical Test Theory

Formulation Mathématique :

$$X = T + e$$

- X = observed score
- T = true score
- e = error

Hypothèses de départ :

- E(X) (l'espérance ou means en anglais) est indépendante de l'erreur si l'échantillon est suffisamment grand.

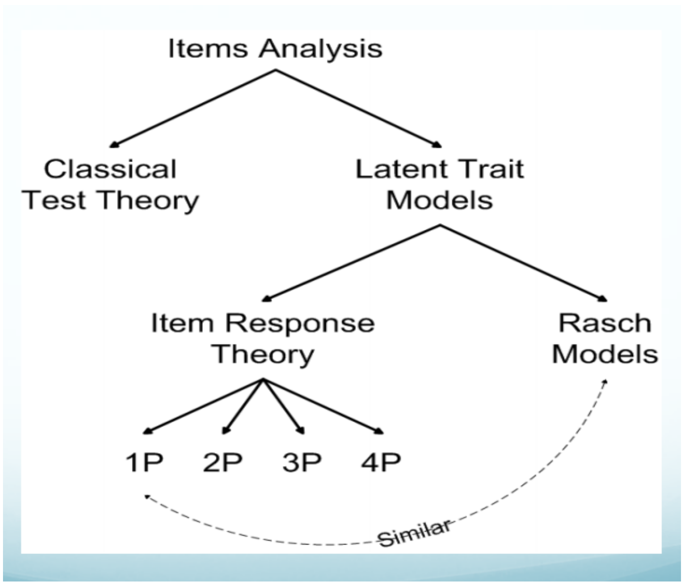
- L'erreur est indépendante item par item.

- e et T sont indépendants (connaître T ne permet pas de déterminer e sans connaître X et vice-versa)

Limite du modèle :

- Une seule valeur de fiabilité (minimisation de l'erreur est le seul moyen de vérifier que le modèle est correct).

- Les scores sont dépendant du test et des personnes qui le passent (on ne peut dire “cette personne à eu 18 à ce test d'histoire donc il aura surement 18 à cet autre test d'histoire”)



<u>Etude de la TRI</u>

<u>Pourquoi la TRI alors plutôt que la CTT ?</u>

La TRI possède non pas une équation globale à l'entièreté du test mais une pour chaque item donc on passe d'un test score model à un item score model, en plus de mesurer la difficulté du test item par item et non pas via une moyenne sur l'ensemble des questions.

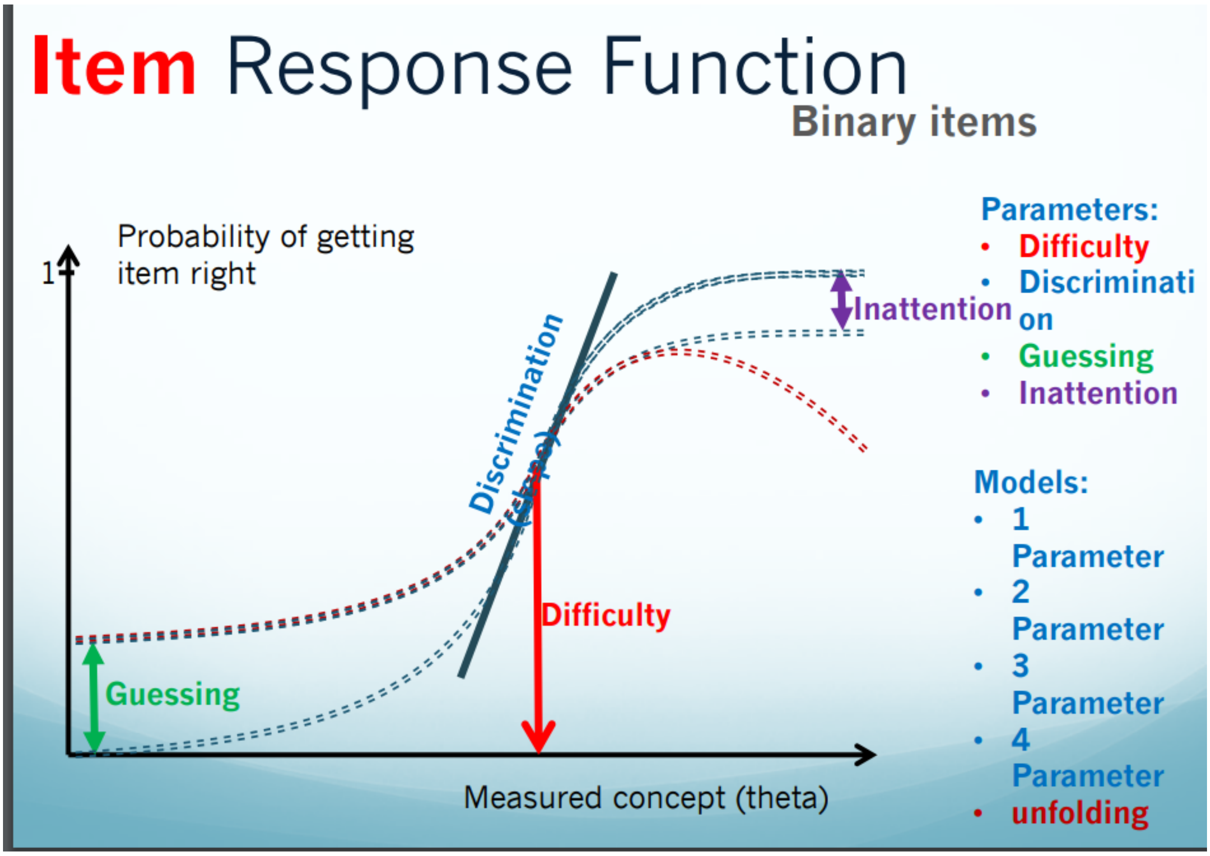
	Classical	**TRI**
Modélisation/Interprétation	Score total	Score items par items
Précision/Information	Idem pour tous les participants/scores	Différent pour chaque participants/scores
Adaptabilité	Impossible	Possible
Score	Dépend de l'ensemble des items	Dépend d'un item
Paramètre des items	Dépend de l'échantillon	Dépend de l'échantillon
Items Privilégiés	Difficulté Moyenne	Toute difficulté

<u>Hypothèses de départ :</u>

- Les items et les personnes passant le test sont indépendants entre eux (On considère que l'ability de chaque participant ne varie pas d'item à item/local indépendance).

- Les items ne mesurent qu'une chose (unidimensionnalité)

<u>Les Composantes de la TRI :</u>



<u>Les mathématiques :</u>

$$p_i(\theta) = c_i + \frac{1 - c_i}{1 + e^{-a_i(\theta - b_i)}}$$

est la formule (que l'on appellera A pour le reste) pour le modèle à trois paramètres d'une distribution normale, qui se trouvent être ai, bi et ci que l'on détaillera. Cette formule peut se simplifier sous la forme $P(t) = \frac{1}{1+e^{-t}}.$ pour permettre des calculs à la main plus facile.

bi, la difficulté : On peut simplifier A au point d'inflexion ce qui nous donne cette formule p(b)=(1+c)/2 .
En effet la probabilité d'avoir bon va osciller entre 1 (tout le monde a bon tout le temps) et ci(qui désigne le minimum de chance d'avoir bon en répondant au hasard).
Elle désigne la hauteur du point d'inflexion de la courbe (en rouge sur le graphique)

ai, la discrimination : La discrimination nous donne l'équation de la dérivé au point d'inflexion, ce qui revient à donner l'allure de la courbe dans cette région (l'équation de la tangente).

La tangente a pour équation p'(b)= a(1-c)/4

ci, la pseudo chance : Elle désigne la probabilité de répondre correctement en s'en remettant totalement au hasard (pour un QCM à 4 choix elle est donc égale à ¼ =0.25).

Exemple : En prenant c=0 , on obtient une hauteur de courbe égale à ½ et une tangente à a/4. Et en injectant cela dans A on obtient que la chance de répondre juste est p(théta)=a(théta -b) ce qui est cohérent car plus la question est difficile plus la probabilité baisse.

Le modèle de Rasch : La grande différence avec le 1PL est que le modèle de Rasch est que le modèle prévaut sur les données, à l'inverse du 1PL qui place les données au-dessus du modèle. En pratique, le modèle de Rasch va mieux fonctionner en moyenne tandis que le 1PL fonctionnera mieux dans les cas extrêmes (modèle orienté).

Vérification du modèle : La méthode Chi-square permet de vérifier une hypothèse (un modèle) en le comparant avec les données obtenues.

The value of the test-statistic is

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = N \sum_{i=1}^{n} \frac{(O_i/N - p_i)^2}{p_i}$$

where

χ^2 = Pearson's cumulative test statistic, which asymptotically approaches a χ^2 distribution.
O_i = the number of observations of type *i*.
N = total number of observations
$E_i = Np_i$ = the expected (theoretical) count of type *i*, asserted by the null hypothesis that the fraction of type *i* in the population is p_i
n = the number of cells in the table.

On compare ensuite la valeur obtenue avec la table de Chi (dispo sur Internet) et on valide ou non l'hypothèse.

<u>L'information :</u>

Tout d'abord selon la théorie de l'information de Fisher la courbe d'information de 1PL suit simplement la fonction qui est la résultante de la multiplication de la probabilité d'une bonne réponse par la probabilité d'une mauvaise réponse. Ceci s'écrit :

$$I(\theta) = p_i(\theta) q_i(\theta).$$

avec pi=correcte et qi=incorrecte

De même l'erreur standard dans l'estimation est :

$$\mathrm{SE}(\theta) = \frac{1}{\sqrt{I(\theta)}}.$$

Pour 2PL et 3PL l'information suit une fonction ou apparaissent en paramètre la détermination et la pseudo chance, soit :

$$I(\theta) = a_i^2 p_i(\theta) q_i(\theta).$$ pour le 2PL et $$I(\theta) = a_i^2 \frac{(p_i(\theta) - c_i)^2}{(1-c_i)^2} \frac{q_i(\theta)}{p_i(\theta)}$$ pour le 3PL