

Prédiction de performance sur des questions dichotomiques : comparaison de modèles pour des tests adaptatifs à grande échelle

Jill-Jênn Vie

EAEI 2015

2 juin 2015

Que sont les tests adaptatifs ?

Des tests où les questions posées dépendent des réponses aux questions précédentes.

Quel est l'intérêt ?

- Obtenir un test plus personnalisé
- Poser des questions comme un humain le ferait
- Profiter du fait que beaucoup de gens ont déjà passé un certain test pour profiler un nouveau venu plus efficacement

- Le choix de la question suivante
(par exemple, la plus discriminante)
- Un critère de terminaison
(dès que l'intervalle de confiance est assez serré ; ou bien dès qu'on a posé 10 questions)

Où pourraient-ils intervenir ?

- Test de positionnement au début d'un cours (par ex. un MOOC)
- Diagnostic efficace des lacunes

Divers types de réponse

- Vrai/Faux
- QCM
- Réponse libre
- Construction de réponse (ex. Castor, via une interface Web)

Nombre :

42

Nombre d'étapes effectuées : 8

+1 x2

Recommencer

Pour nos expériences, nous avons à disposition :

- un test américain multidisciplinaire : le **SAT**
(300 étudiants, 20 questions)
- un test de mathématiques sur les **fractions**
(500 étudiants, 20 questions)
- le concours d'informatique **Castor**
(60000 étudiants 6e-5e, 17 questions)

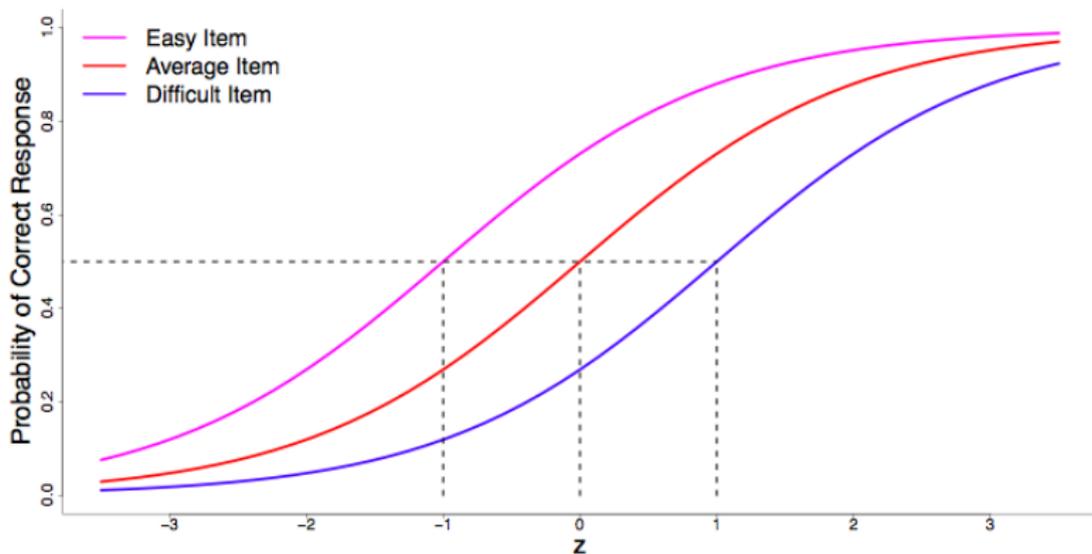
Le format des données étudiant

	Q_1	Q_2	Q_3	\dots	Q_{19}	Q_{20}
Alice	F	V	V	\dots	F	V
Bob	V	F	V	\dots	V	F
Charles	F	F	F	\dots	F	F
\vdots	\vdots	\vdots	\vdots	\dots	\vdots	\vdots

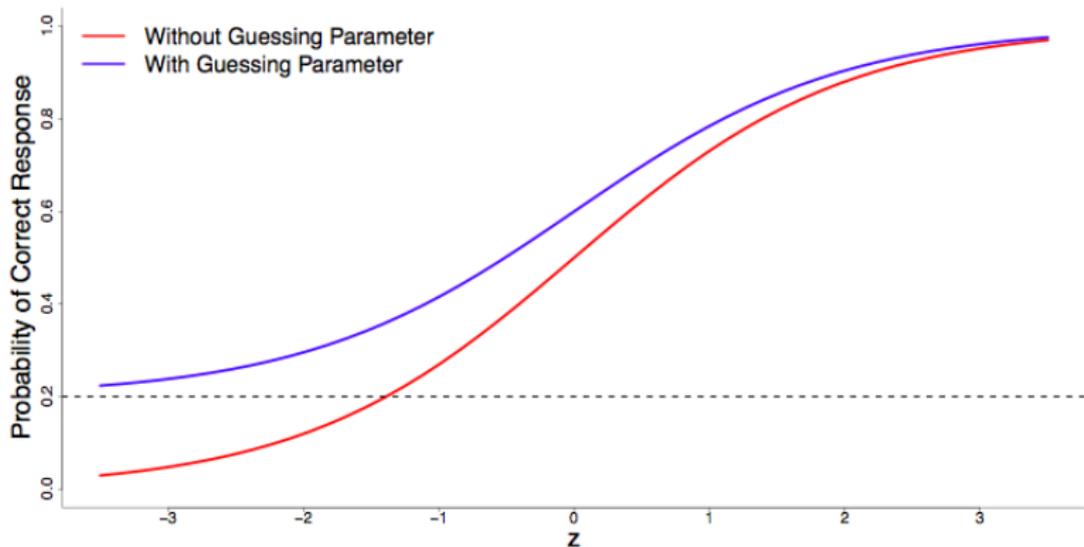
Chaque ligne est appelée **vecteur de réponse**.

- On part d'un ensemble de questions qui ont été posées à un grand nombre d'apprenants
- De nouveaux apprenants arrivent, on détermine des questions adaptées à leurs connaissances

Le paramètre de difficulté



Le paramètre de **guessing**

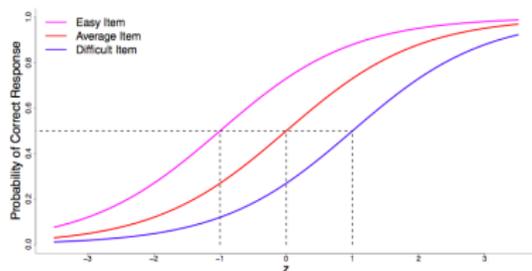


	C_1	C_2	C_3
Q_1	1	0	0
Q_2	0	1	1
Q_3	1	1	0
\vdots	\vdots	\vdots	\vdots

- Elle indique quelles compétences sont **requis** pour répondre correctement à chaque question.
- Toutes compétences maîtrisées \rightarrow probabilité de se tromper
- Pas toutes maîtrisées \rightarrow probabilité de répondre correctement
- Spécifiée par un expert **ou** calculée par un algorithme

La question dont la probabilité de réponse est **le plus proche de $1/2$**
(celle qui maximise l'information de Fisher)

Réponse à l'item



- Difficulté des questions
- Niveau des candidats
- Ordre total entre les candidats

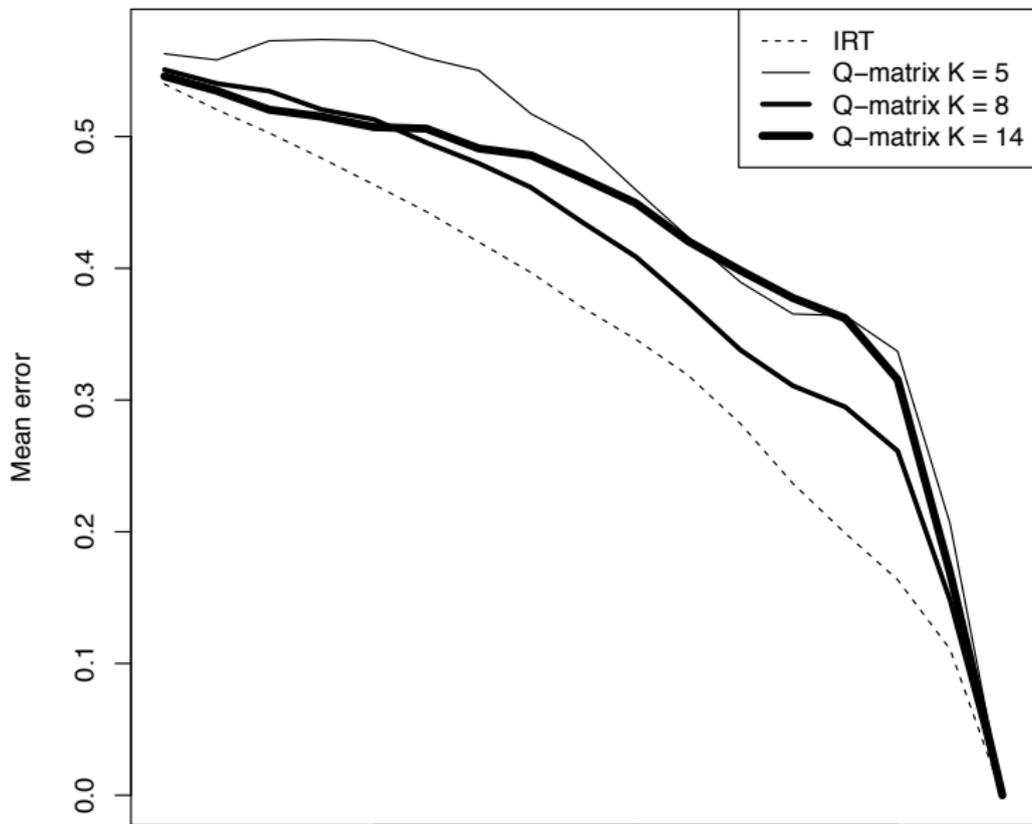
Q-matrice

	C_1	C_2	C_3
Q_1	1	0	0
Q_2	0	1	1
Q_3	1	1	0
\vdots	\vdots	\vdots	\vdots

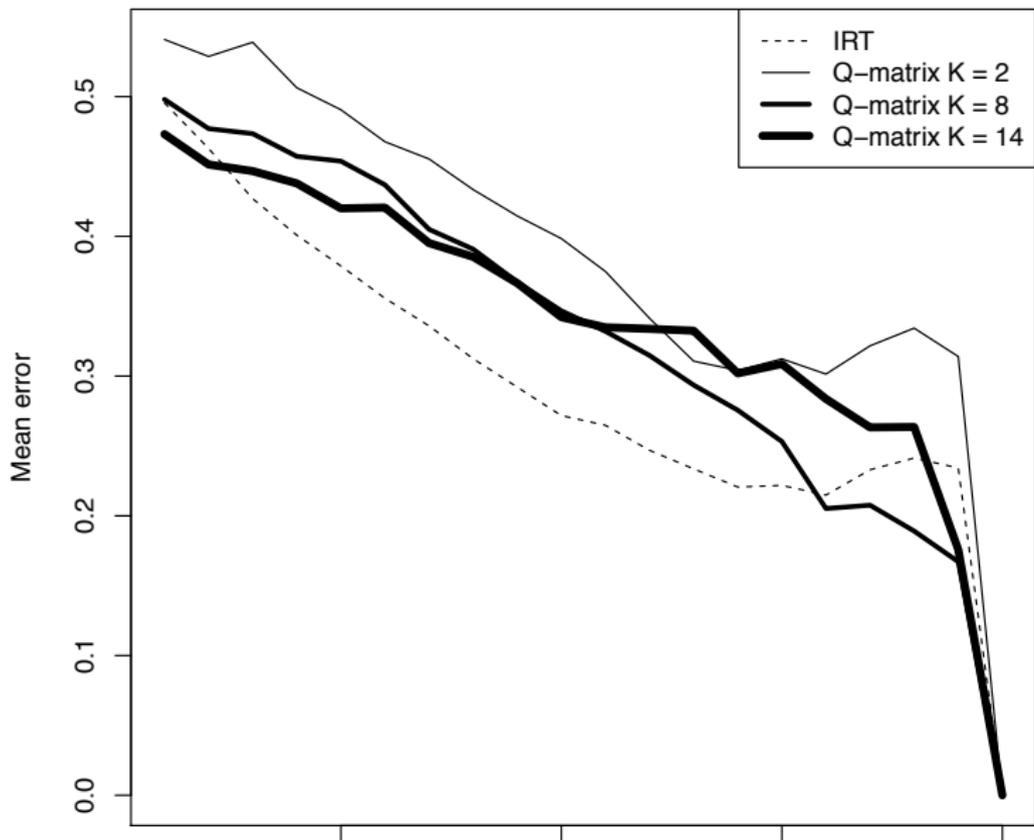
- Compétences requises pour résoudre chaque question
- Maîtrise ou non-maîtrise de chaque compétence par chaque candidat
- Ordre partiel entre les candidats

- ① On **entraîne** les modèles avec 80 % des étudiants.
Pour les 20 % restants → vecteurs de réponse « inconnus ».
- ② Après avoir posé un certain nombre de questions, pour toutes les questions non encore posées :
→ **inférence** que le candidat y réponde correctement.
- ③ On **évalue** les modèles : plus l'estimation de réponse correcte ou incorrecte est loin de la réalité, plus la pénalité est grande.

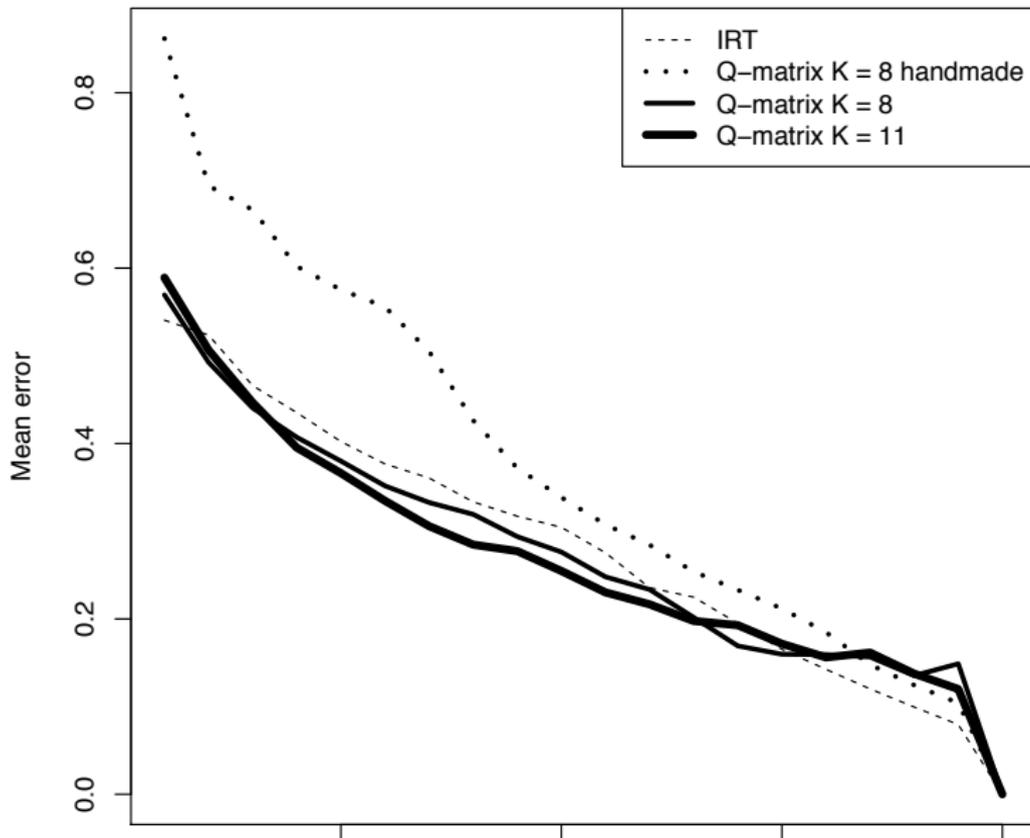
Résultat : Castor



Résultat : SAT



Résultat : Fraction



- Le modèle qui prédit le mieux semble dépendre des caractéristiques du test
- Une q-matrice déterminée par un algorithme a un meilleur pouvoir prédictif mais n'a pas de compétences explicites

- Comparer au modèle **Additive Factors** : plus le niveau d'un candidat est grand et plus il a de compétences, plus sa probabilité de répondre correctement est grande
- Déterminer à quel point la note d'un candidat est corrélée à celle de ses **10 plus proches voisins** (pour la distance sur les vecteurs de réponses)
- Développer un outil de tests adaptatifs en Python pour le relier à Open edX

Vous pouvez m'écrire à jjv@lri.fr

<http://jiji.cat>

<http://jill-jenn.net>