



Modèle à partition ordonnée pour l'analyse de questionnaires avec degrés de certitude

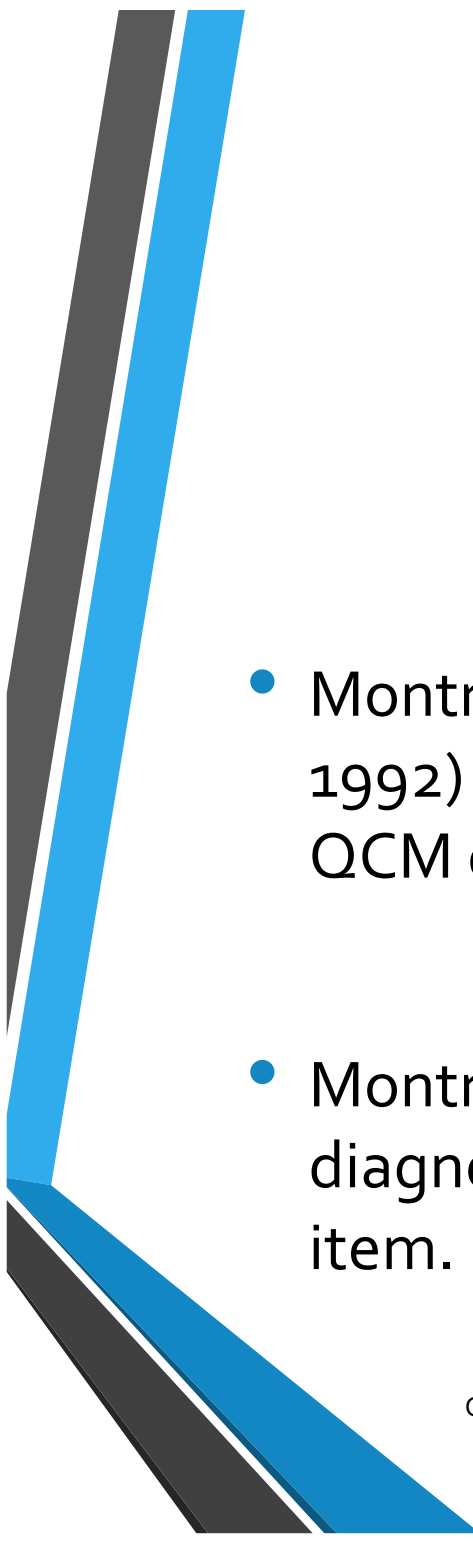
Oliver Prosperi, HEP Vaud et Université de Fribourg



Oliver Prosperi, HEP Vaud, Université de Fribourg.
oliver.prosperi@alumni.epfl.ch



UNIVERSITÉ DE FRIBOURG
UNIVERSITÄT FREIBURG

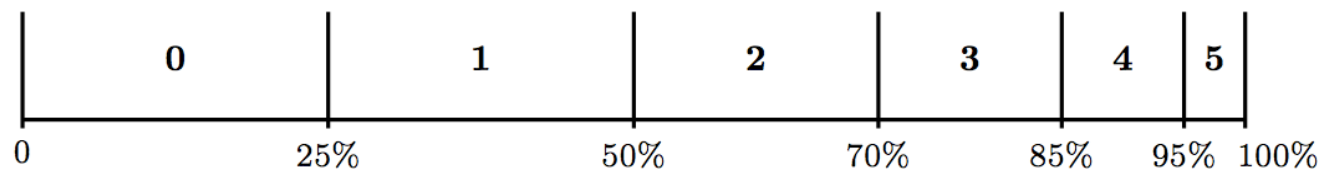


But de la recherche

- Montrer que le Modèle à Partition Ordonnée (Wilson, 1992) peut être utilisé pour modéliser les item de test par QCM employant les Degrés de Certitude
- Montrer que la modélisation peut être utilisé comme outil diagnostique pour détecter des “problèmes” dans un item.

Les Degrés de Certitude (DC)

- **Probabilité subjective de répondre correctement à une question.**
- **Les Intervalles** de degrés de certitudes



- **L'Echelle des scores**

		Degrés de certitude					
		0	1	2	3	4	5
Réponse	Correcte	+13	+16	+17	+18	+19	+20
	Incorrecte	+4	+3	+2	0	-6	-20

Modèle à partition ordonné

- Fonction $B(k)$ *ad hoc* qui attribue un score à chaque catégorie de performance $k = 1, 2, \dots, K$.
- Soit I le nombre d'items $i = 1, 2, \dots, I$ et soit K le nombre de catégories de performance. Soit X_{ni} une variable aléatoire représentant la réponse de l'étudiant n à l'item i et θ_n son niveau d'aptitude. Alors sa probabilité d'être dans la catégorie k pour la question i est donnée par:

$$P(X_{ni} = k) = \frac{\exp[\theta_n B(k) - \xi_{ik}]}{\sum_{h=1}^K \exp[\theta_n B(h) - \xi_{ih}]}$$

- Où ξ_{ik} est le paramètre associé à la catégorie k pour l'item i , et où $\xi_{i1} \equiv 0$.
- Contrainte d'identification: $\sum_{k=1}^K \xi_{ik} = 0$



Modèle à partition ordonné

- Modèle implémenté dans le logiciel ACER ConQuest
- Détails sur l'estimation du modèle et une discussion de son champs d'application dans:

Wilson, M., et Adams, R. J. (1993). Marginal maximum likelihood estimation for the ordered partition model. *Journal of Educational Statistics*, 18(1), 69-90.

Modèle à partition ordonné

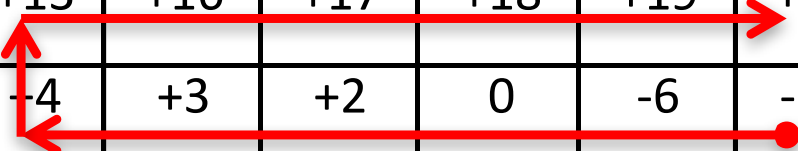
- Extension du PCM
- Si $B(k)=k$ pour tout $k=1,2,\dots,K$, alors le modèle est équivalent au PCM. Par exemple comme dans

Pruski, L.A., Blanco, S.L., Riggs, R.A., et Lichtenstein, M.J. (1993). Construct Validation of the Self-Efficacy Teaching and Knowledge Instrument for Science Teachers-Revised (SETAKIST-R): Lessons Learned. *Journal of Science Teacher Education*, 24(7), 1133-1156

Comment?

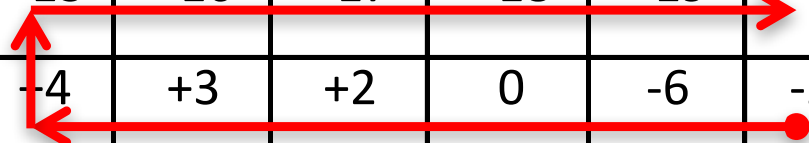
- Pour les DC, K sera le nombre de cases de l'échelle des scores, ainsi $K=12$.
- Les 12 catégories de performance sont définies par :

	Degré de certitude					
	0	1	2	3	4	5
Correcte	+13	+16	+17	+18	+19	+20
Incorrecte	-4	+3	+2	0	-6	-20

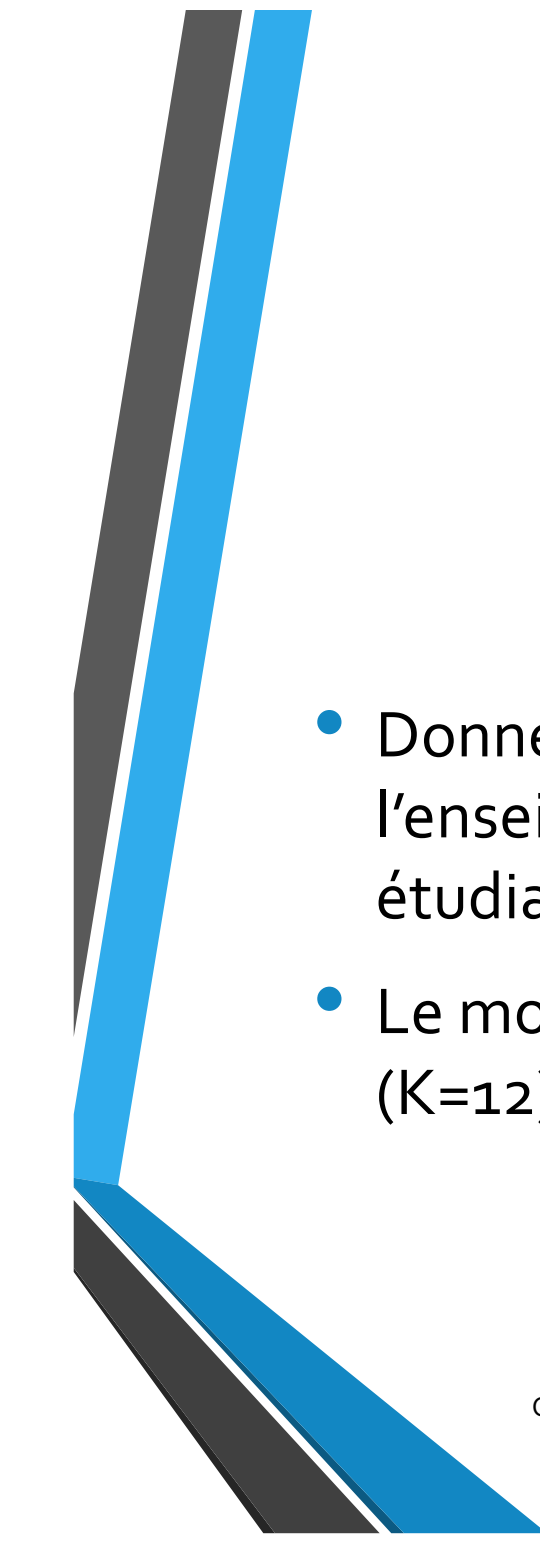


Comment?

	Degré de certitude					
	0	1	2	3	4	5
Correcte	+13	+16	+17	+18	+19	+20
Incorrecte	+4	+3	+2	0	-6	-20



- $\text{Cat1}=\text{DC}_5+\text{RI}$; $\text{Cat2}=\text{DC}_4+\text{RI}$; ... ; $\text{Cat6}=\text{DC}_0+\text{RI}$; $\text{Cat7}=\text{DC}_0+\text{RC}$; ; $\text{Cat12}=\text{DC}_5+\text{RC}$
- $B(1,2,\dots,6)=0$ and $B(7,8,\dots,12)=1$

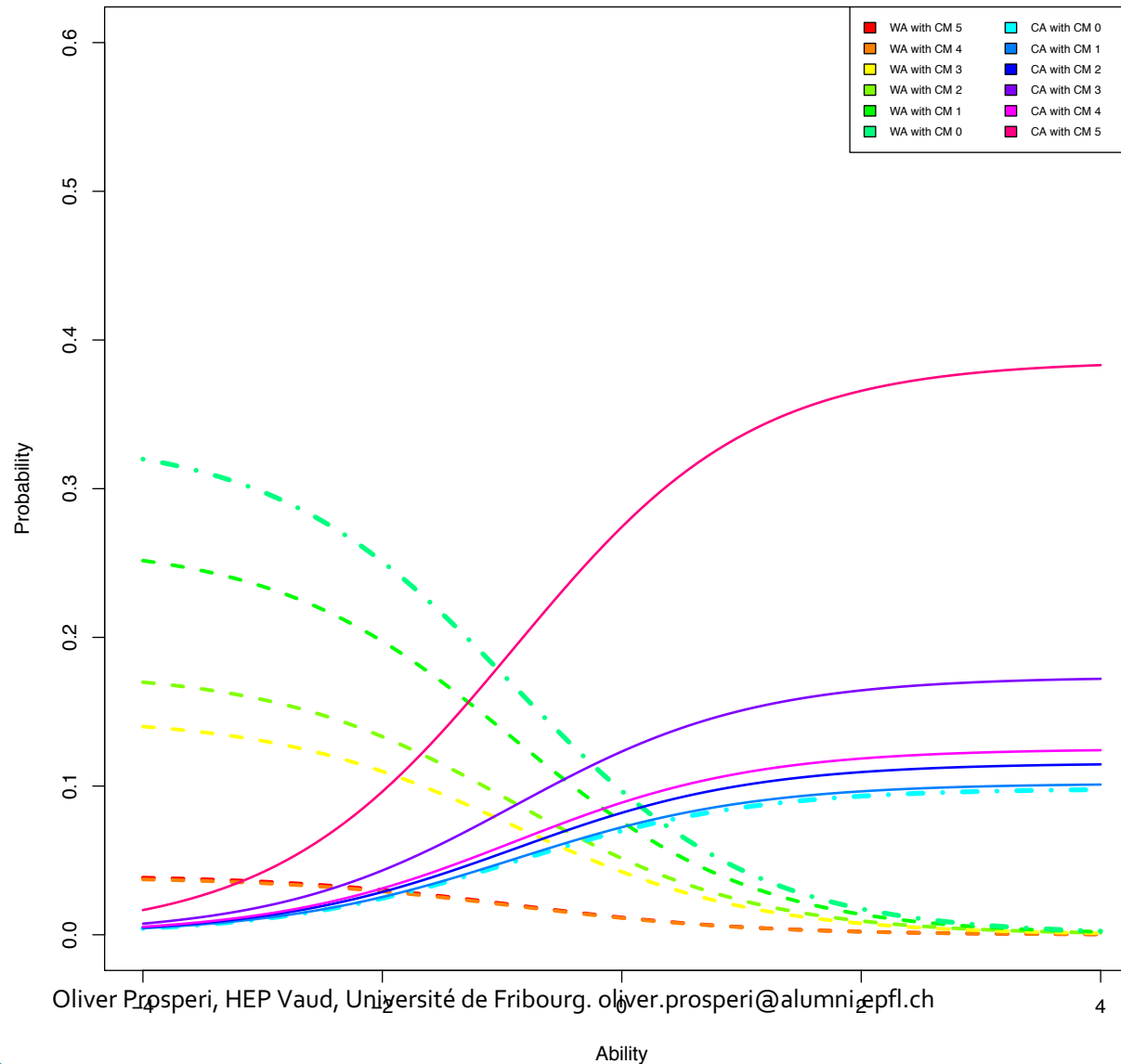


Résultats

- Données du test EFES (Evaluation du français pour l'enseignement supérieur) 2011, 60 items, 2'531 étudiants.
- Le modèle produit un paramètre pour chaque catégorie k ($K=12$) -> Courbes caractéristiques!

Courbes Caractéristiques

Item 7

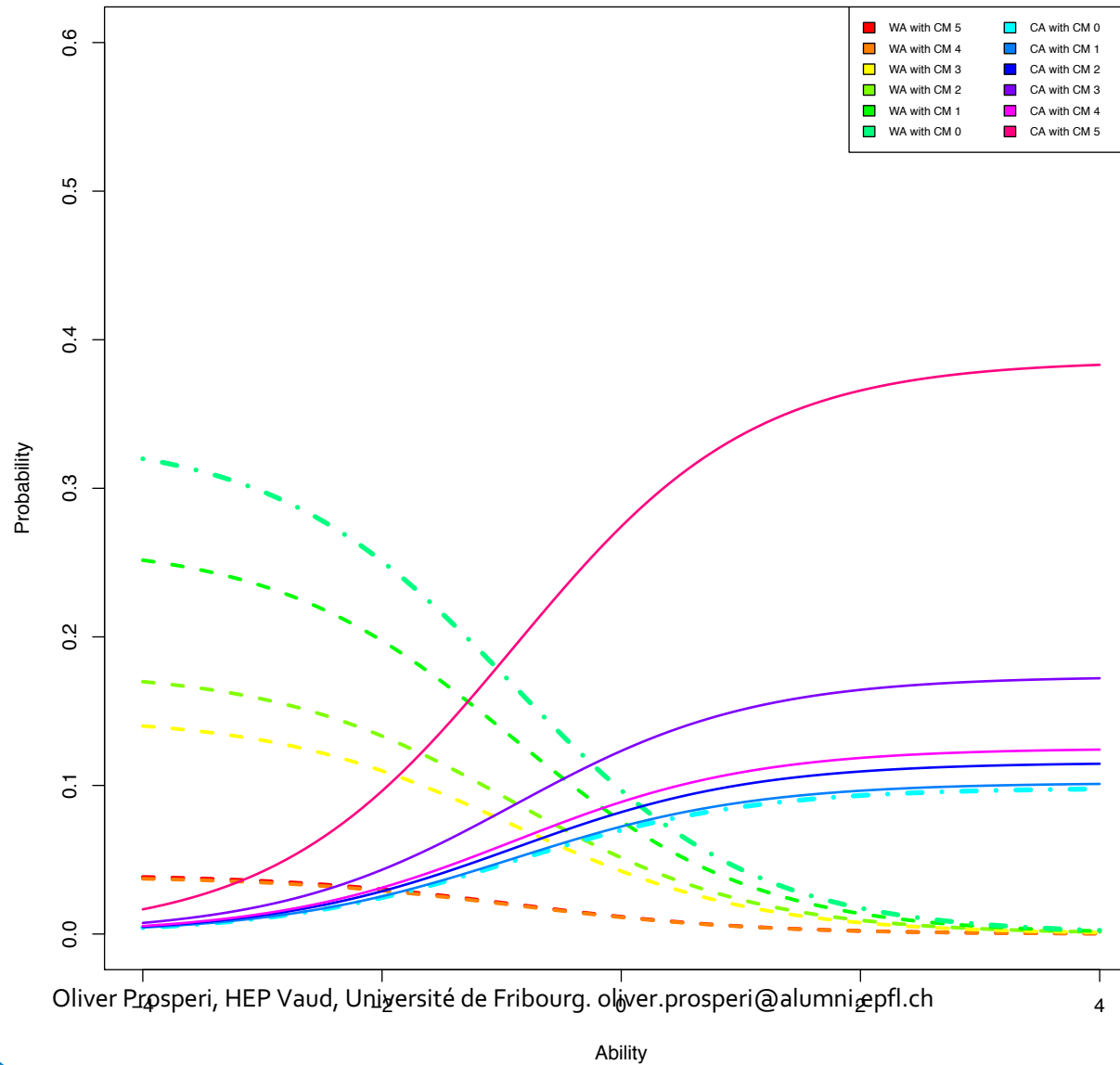


Résultats

- Données du test EFES (Evaluation du français pour l'enseignement supérieur) 2011, 60 items, 2'531 étudiants.
- Le modèle produit un paramètre pour chaque catégorie k ($K=12$) -> Courbes caractéristiques!
- Eclatement de la courbe de Rasch

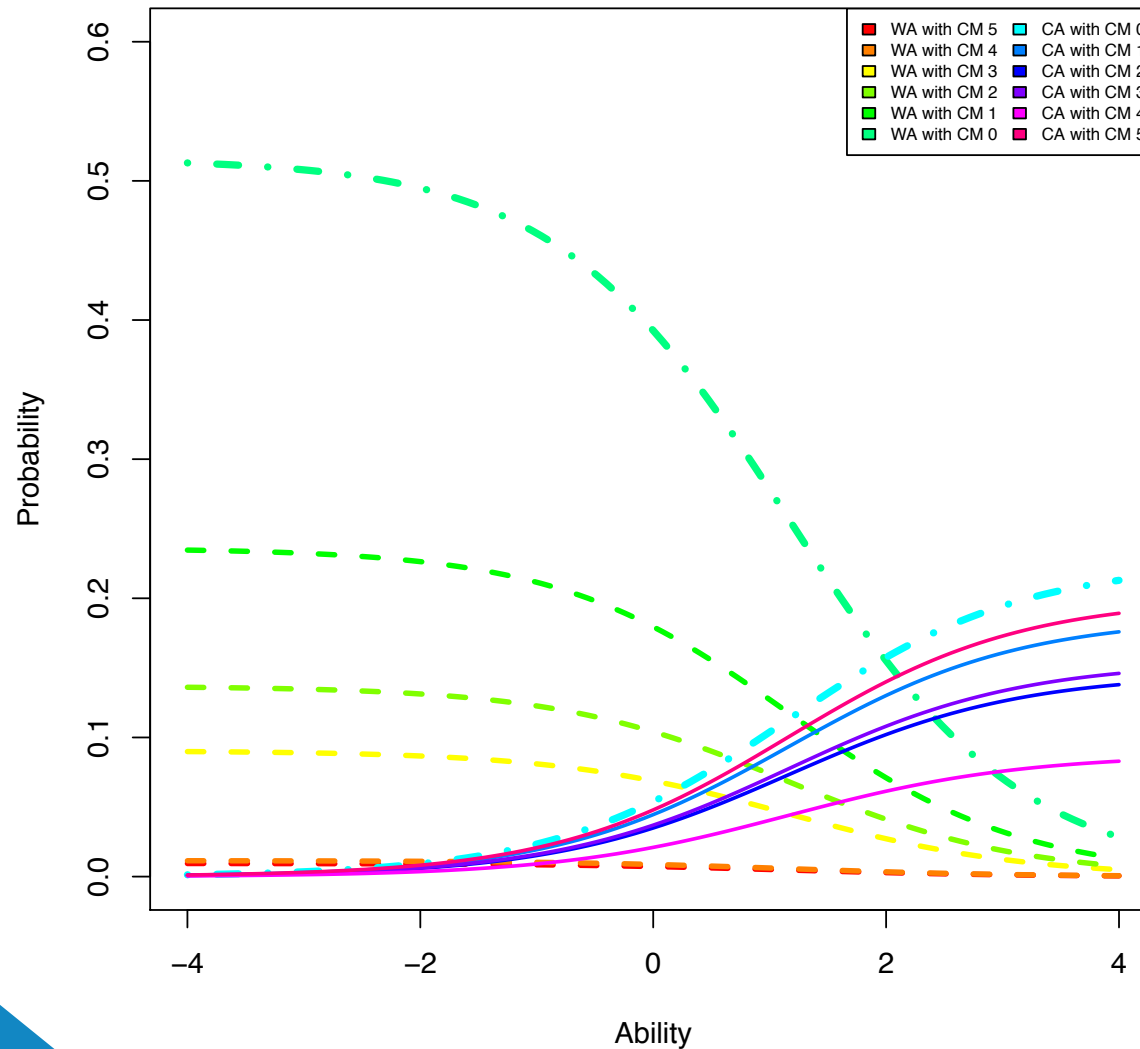
Courbes Caractéristiques

Item 7



Courbes Caractéristiques

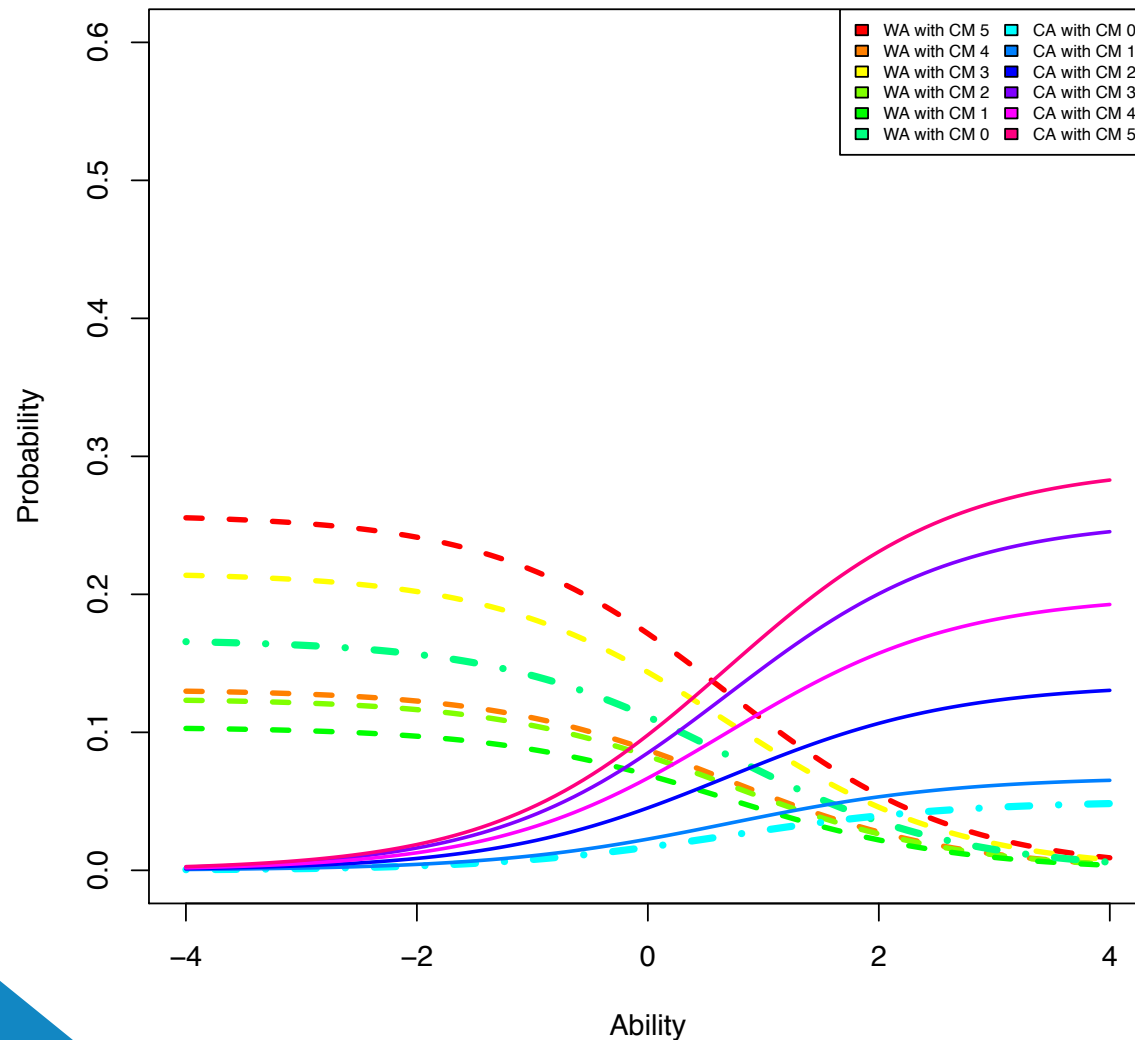
Item 12



Les deux courbes supérieures sont celles du DCo : **les étudiants ont probablement essayé de deviner, même quand l'aptitude monte. Les étudiant très forts répondent néanmoins correctement**

Courbes Caractéristiques

Item 50



La courbe supérieure pour une réponse incorrecte est celle du DC5: les **étudiants avec une faible aptitude ont surestimé leur compétence** et ont donné une réponse incorrecte avec DC5. Les bons étudiants ont répondu correctement. **L'item pourrait être une question piège!**

Simulation de données DC

- Génération de « bons » paramètres pour simuler des données avec DC:
- La chance de se trouver dans la catégorie k par rapport à la catégorie k' (Wilson 1992) est $\frac{e^{\xi_{ik}}}{e^{\xi_{ik'}}$.
- Avec des contraintes de bord liés aux intervalles de certitude et la contrainte d'identification on peut trouver des sous ensembles de $\mathbb{R}^K$ de « bons » paramètres.

Simulation de données DC

- Exemple avec 2 degrés de certitude: $k \in \{1,2,3,4\}$
- Intervalles de certitude (0%-60%) et (60%-100%)
- Système de contraintes avec 4 inconnues (x_1 pour ξ_{i1} ; x_2 pour ξ_{i2} ; etc.) :

$$\begin{cases} e^{x_4} / e^{x_1} \geq 60/40 \\ e^{x_3} / e^{x_2} < 60/40 \\ x_1 + x_2 + x_3 + x_4 = 0 \end{cases}$$



Simulation de données DC

- Grapher!



Merci!

Contact:
Oliver.prosperi@alumni.epfl.ch