

*La détection des cas de fraudes lors  
de l'évaluation des apprentissages et  
une approche pour corriger les  
résultats malgré le comportement  
frauduleux*

**Gilles Raïche, UQAM**

**12 novembre 2009**

**Pavillon des sciences de l'éducation**

**12 h 30 à 14 h 00 - Local N-4630**

Université du Québec à Montréal

Centre sur les applications des modèles de réponses aux items (CAMRI)

<http://www.camri.uqam.ca/camri/camriBase/>



# PLAN DE LA PRÉSENTATION

- Intérêts de recherche et buts visés
- Causes des patrons de réponses inappropriés
- Score à un test et indices de détection
- Modélisations de l'habileté
- Modélisations de l'ajustement
- Modélisations multidimensionnelles alternatives
- Application à la correction de la mesure de l'habileté
- Autres applications
- Application logicielle: *irtProb*
- Références





## INTÉRÊTS ET BUTS

- **Détection des patrons de réponses inappropriés**
- **Correction de la mesure de l'habileté ou du trait malgré l'inadéquation du patron de réponses**
- **Création d'une modélisation de la réponse à un item intégrant des paramètres de sujet autres que son habileté uniquement**
- **Modélisation permettant de simuler des patrons de réponses inappropriés**
- **Élaboration de meilleures modélisations cognitives de la réponses à un item**



# CAUSES DES PATRONS DE RÉPONSES INAPPROPRIÉS

- **Plagiat**
- **Sous performance intentionnelle**
- **Difficulté de compréhension des questions ou des réponses**
- **Inattention**
- **Réponses au hasard**
- **Etc.**
  
- **Globalement :**
  - *Response set* (Cronbach, 1950)
  - *Response style* (Frederiksen et Messick, 1956)
  - *Biased response* (Berg, 1955)
  - *Stylistic variance* (Wiggins, 1962)



# SCORES À UN TEST ET INDICES DE DÉTECTION

- **Score à un test (*content set*)**
- **Score d'ajustement (*response set*)**
- **Indices de détection (*person fit index*)**



# MODÉLISATION DE L'HABILETÉ (1)

- **Théorie classique des tests**

$$SCORE\ VRAI_j = TOTAL_j + ERREUR_j$$

- **Ce modèle implique :**
  - Une relation linéaire entre la réponse à un item et le score au test
  - L'erreur est la même quelle soit le score au test
  - Une théorie des tests et non pas une théorie des réponses aux items
  - Ne permet pas de simuler de façon réaliste des réponses à des items



## MODÉLISATION DE L'HABILETÉ (2)

- **Théorie de la réponse à l'item (modèle général)**

La probabilité qu'un étudiant  $j$  obtienne une bonne réponse ( $x_{ij}=1$ ) à un item  $i$  dépend du niveau d'habileté (*thêta*) de l'étudiant et des paramètres de l'item (*zeta*)

$$P(x_{ij} = 1 | \theta_j, \zeta_i)$$

La probabilité qu'un étudiant  $j$  obtienne un patron de réponses correspond au produit suivant :

$$P(X_j | \theta_j, \zeta) = \prod_i \left\{ P(x_{ij} = 1 | \theta_j, \zeta_i)^{x_{ij}} P(x_{ij} = 0 | \theta_j, \zeta_i)^{1-x_{ij}} \right\}$$



# MODÉLISATION DE L'HABILETÉ (3)

## • Théorie de la réponse à l'item

Modèles logistiques à 1, 2, 3 et 4 paramètres

- a – paramètre de discrimination
- b – paramètre de difficulté
- c – paramètre de pseudo chance
- d – paramètre d'inattention

$$P(x_{ij} = 1 | \theta_j, b_i) = \frac{1}{1 + e^{-(\theta_j - b_i)}} \quad P(x_{ij} = 1 | \theta_j, a_i, b_i) = \frac{1}{1 + e^{-a_i(\theta_j - b_i)}}$$

$$P(x_{ij} = 1 | \theta_j, a_i, b_i, c_i) = c_i + \frac{1 - c_i}{1 + e^{-a_i(\theta_j - b_i)}}$$

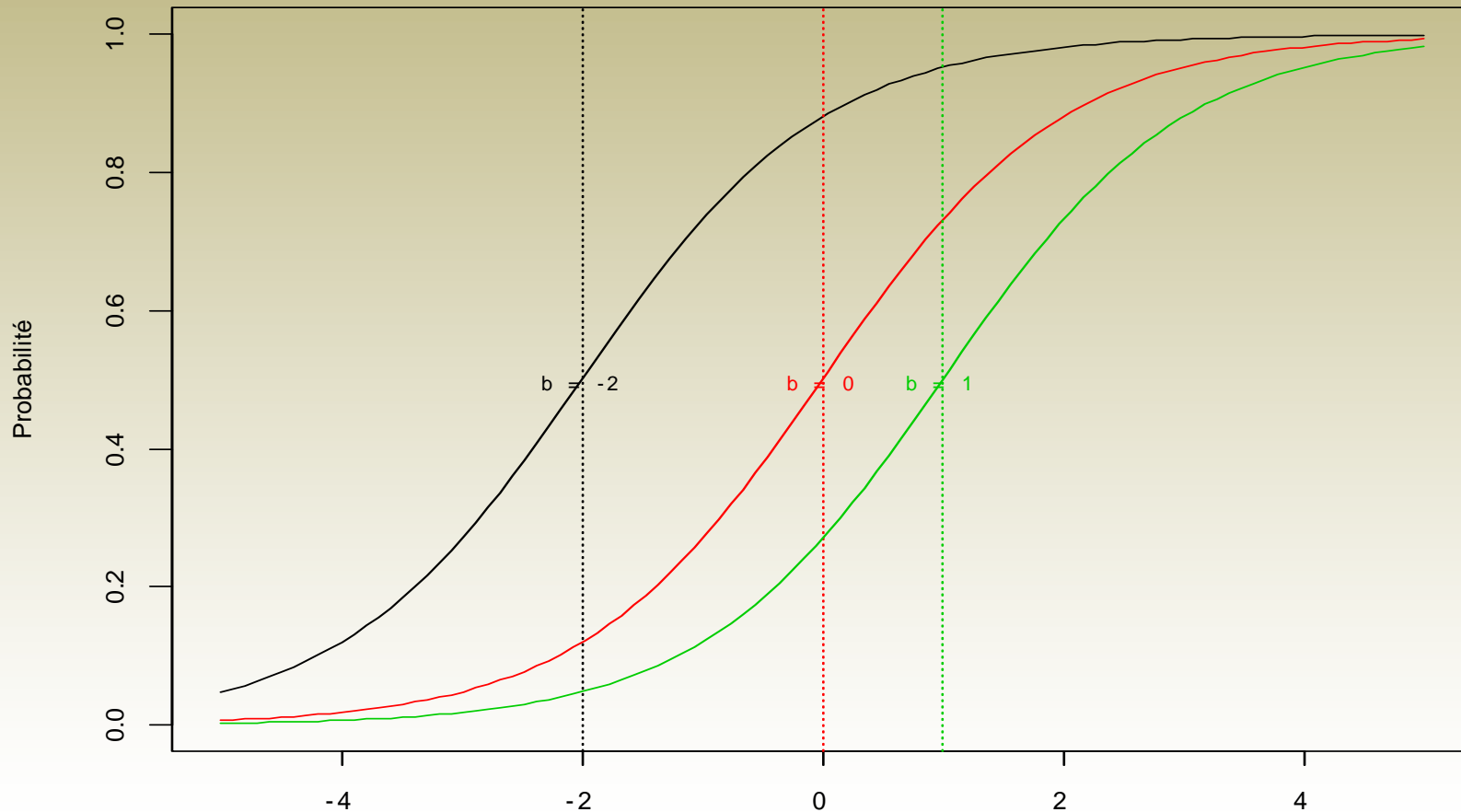
$$P(x_{ij} = 1 | \theta_j, a_i, b_i, c_i, d_i) = c_i + \frac{d_i - c_i}{1 + e^{-a_i(\theta_j - b_i)}}$$





# MODÉLISATION DE L'HABILETÉ (4)

- Théorie de la réponse à l'item (modèle à 1 paramètre)

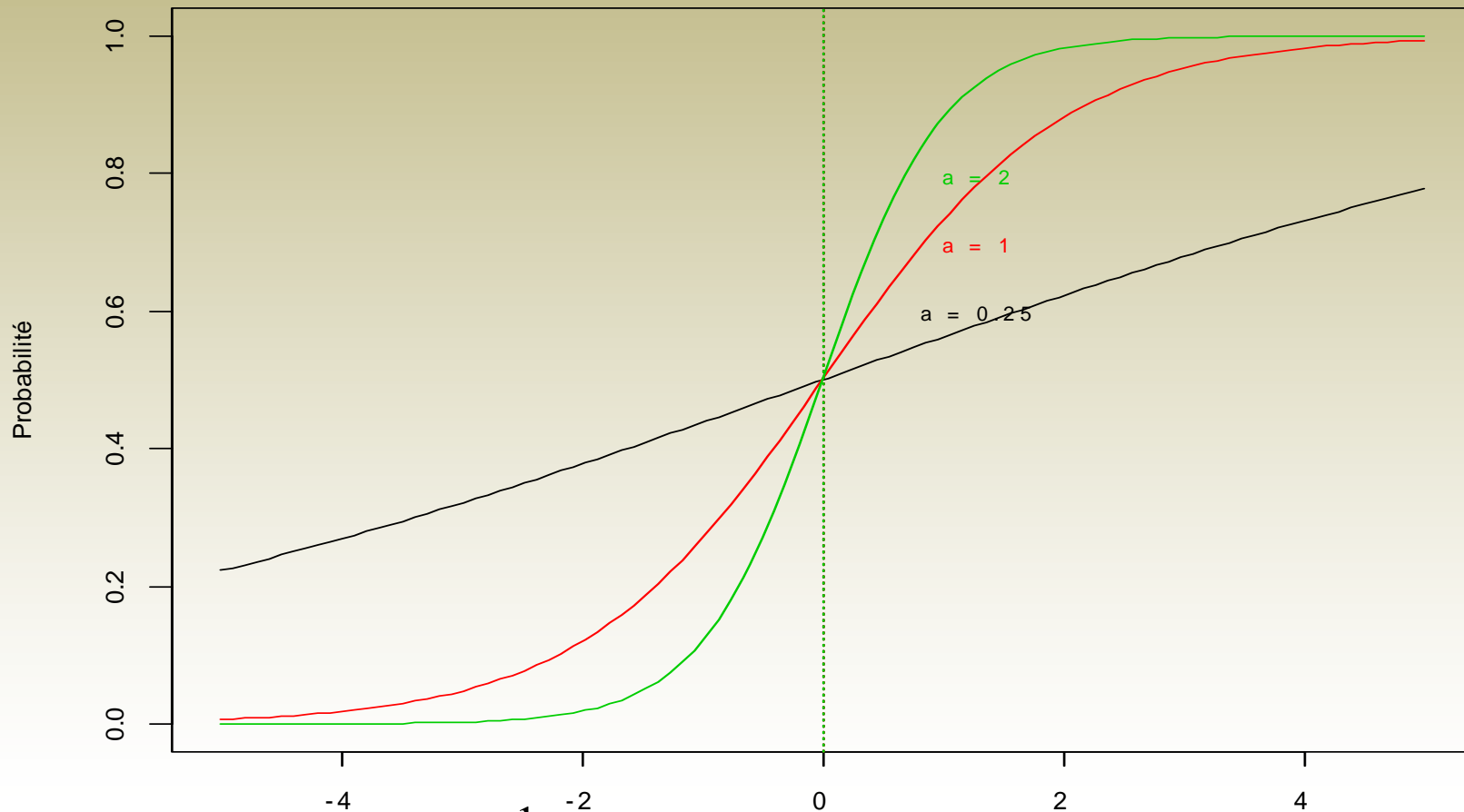


$$P(x_{ij} = 1 | \theta_j, b_i) = \frac{1}{1 + e^{-(\theta_j - b_i)}}$$



# MODÉLISATION DE L'HABILETÉ (5)

- Théorie de la réponse à l'item (modèle à 2 paramètres)

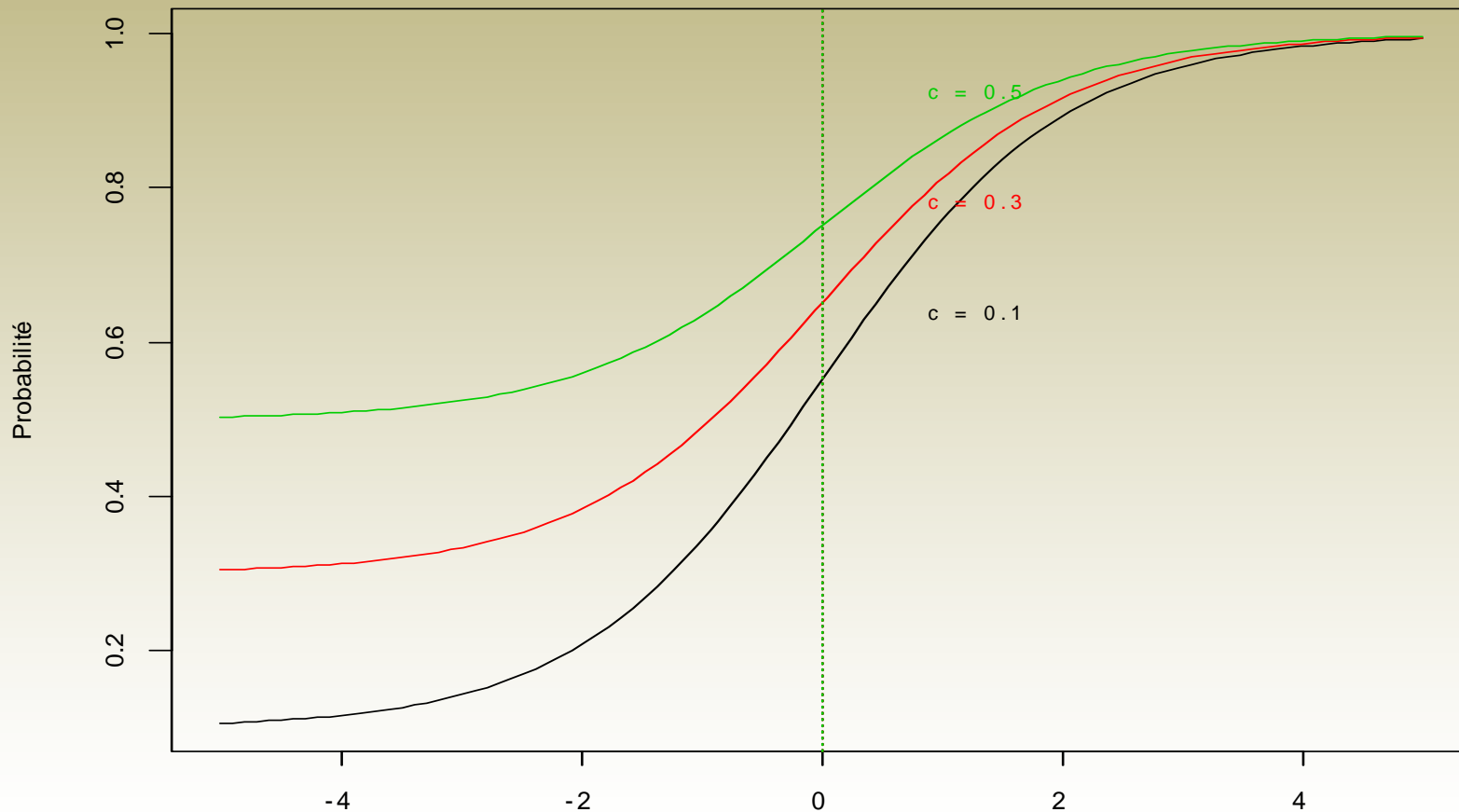


$$P(x_{ij} = 1 | \theta_j, a_i, b_i) = \frac{1}{1 + e^{-a_i(\theta_j - b_i)}}$$



# MODÉLISATION DE L'HABILETÉ (6)

- Théorie de la réponse à l'item (modèle à 3 paramètres)

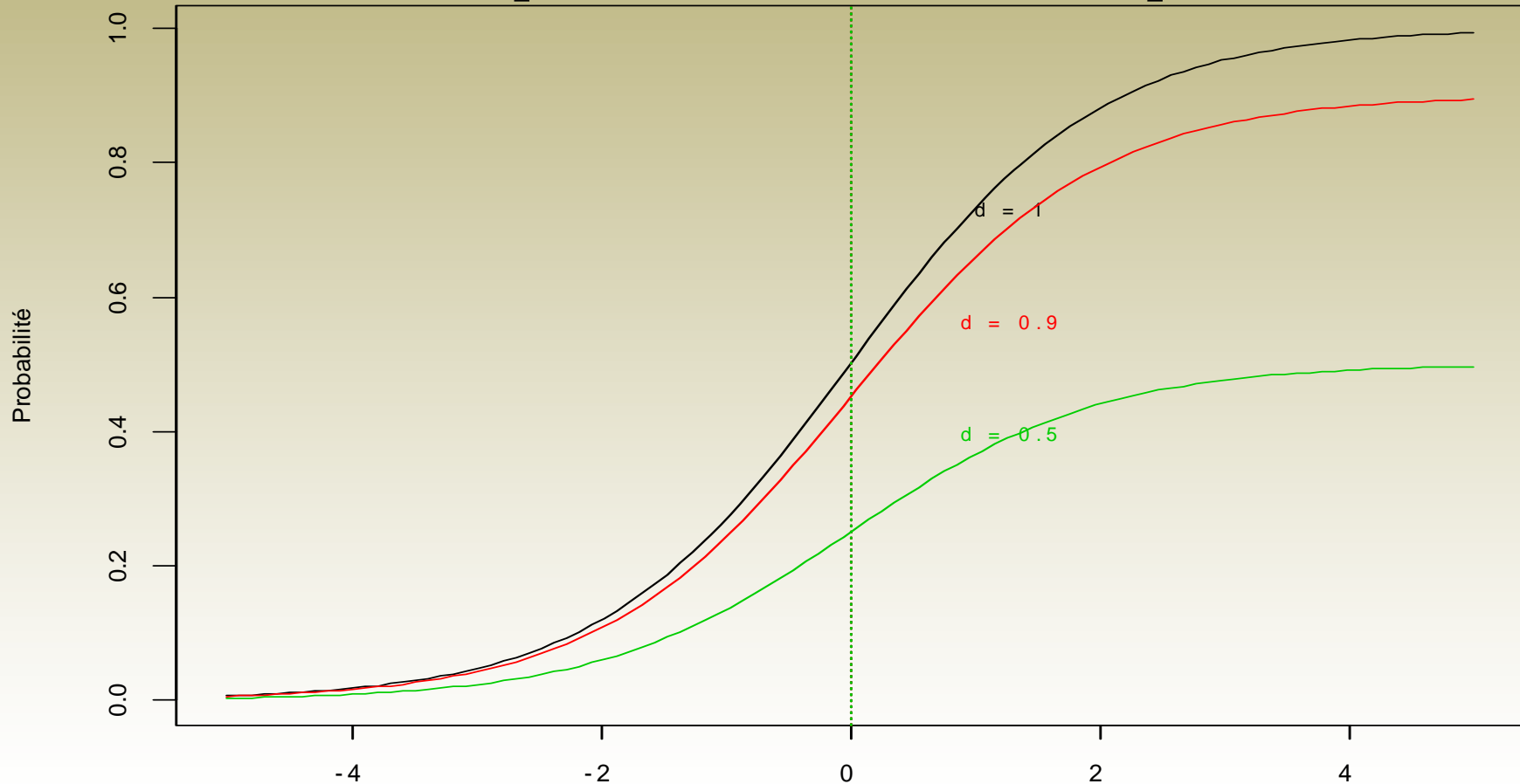


$$P(x_{ij} = 1 | \theta_j, a_i, b_i, c_i) = c_i + \frac{1 - c_i}{1 + e^{-a_i(\theta_j - b_i)}}$$



# MODÉLISATION DE L'HABILETÉ (7)

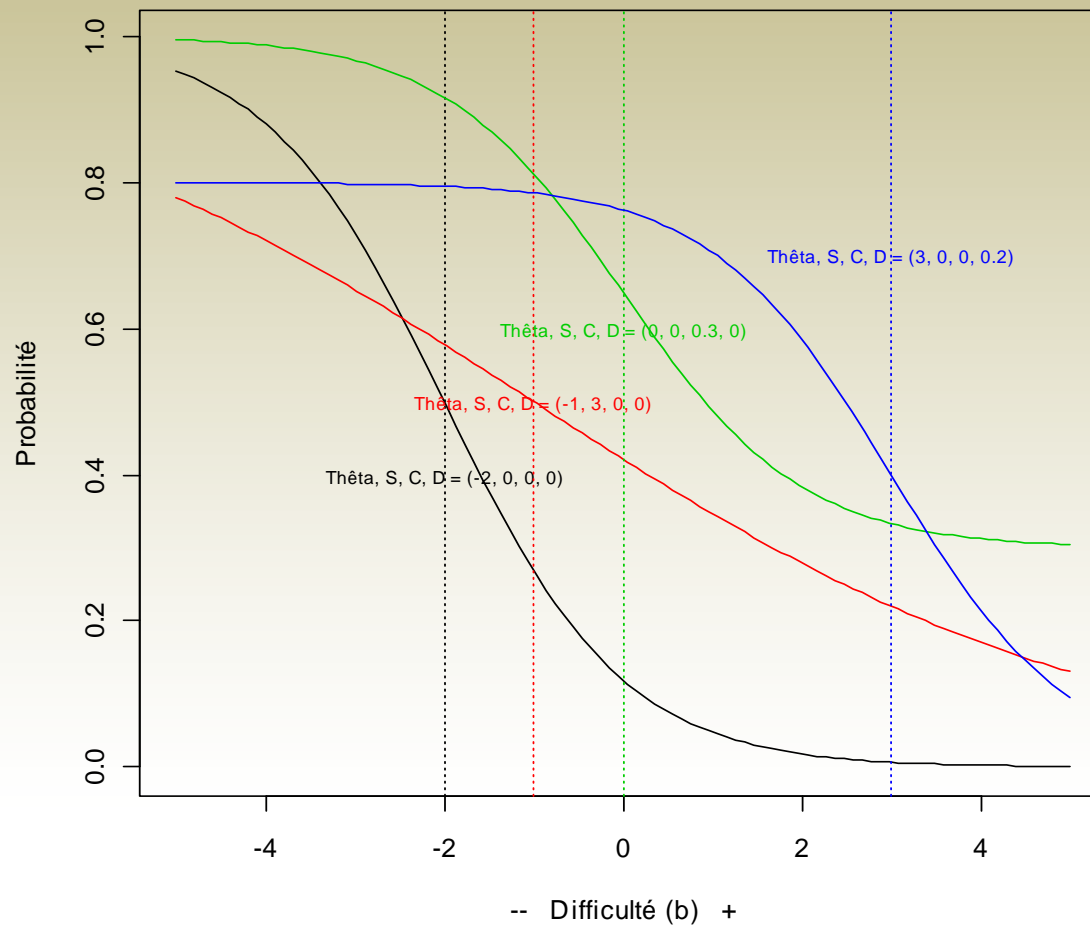
- Théorie de la réponse à l'item (modèle à 4 paramètres)



$$P(x_{ij} = 1 | \theta_j, a_i, b_i, c_i, d_i) = c_i + \frac{d_i - c_i}{1 + e^{-a_i(\theta_j - b_i)}}$$

# COURBE CARACTÉRISTIQUE DE PERSONNE

- Théorie de la réponse à l'item  
(courbe caractéristique de personne)





# MODÉLISATIONS MULTIDIMENSIONNELLES ALTERNATIVES

- **Théorie de la réponse à l'item (modèle multidimensionnel)**

Modèle auquel on ajoute trois paramètres de sujet additionnels :  $S_j$ ,  $C_j$  et  $D_j$ . Ces paramètres sont les équivalents sujets des mêmes paramètres d'items et se distinguent ici par la majuscule.

$$P(x_{ij} = 1 | \theta_j, S_j, C_j, D_j, a_i, b_i, c_i, d_i) = (C_j + c_i) + \frac{(d_i - D_j) - (C_j + c_i)}{1 + e^{-a_i(\theta_j - b_i)}}$$

Noter que  $a_i$  correspondait en fait à l'inverse de l'écart type du paramètre de difficulté ( $1/\sqrt{s_i^2}$ ). Maintenant l'écart type provient à la fois de l'item et à la fois du sujet  $1/(\sqrt{S_j^2 + s_i^2})$ .

$$a_i(\theta_j - b_i) = \frac{-(\theta_j - b_i)}{\sqrt{S_j^2 + s_i^2}}$$

$$P(x_{ij} = 1 | \theta_j, S_j, C_j, D_j, s_i, b_i, c_i, d_i) = (C_j + c_i) + \frac{(d_i - D_j) - (C_j + c_i)}{1 + e^{\frac{-(\theta_j - b_i)}{\sqrt{S_j^2 + s_i^2}}}}$$



# APPLICATION À LA CORRECTION DU NIVEAU D'HABILITÉ (1)

## **Données simulées**

100 sujets simulés pour chacune des conditions

$\theta = -2$  et  $2$

$C = 0,00, 0,10, 0,20, 0,30$  ou  $0,40$

$S = 0,00, 0,50, 1,00, 2,00$  ou  $4,00$

$D = 0$

## **Test administré**

40 items

$b = -5$  à  $5$ ,  $c = 0$ ,  $d = 0$

## **Modèles étudiés**

Modèle 1 :  $\theta$  uniquement

Modèle 2 :  $\theta$  et  $C$

Modèle 3 :  $\theta$ ,  $S$ ,  $C$  et  $D$

## **Méthode d'analyse**

Calcul des estimateurs et de leur écart type empirique (précision)

Méthode d'estimation : MAP (Nelder et Mead) avec une distribution a priori uniforme  $U(-3,3)$  non informative pour le niveau d'habileté seulement

Comparaison sous forme de tableau de ceux-ci selon différentes modèles



# APPLICATION À LA CORRECTION DU NIVEAU D'HABILITÉ (2)

**Tableau 1. Effet de l'utilisation des nouvelles modélisations sur la correction de l'estimation du niveau d'habileté de 100 sujets qui cherchent à augmenter leur résultat à un test de 40 items**

		C (paramètre de pseudo chance personnelle)				
Modèle <sup>1</sup>		0,00	0,10	0,20	0,30	0,40
1	$\theta = -2$	-2,18 (0,62)	-1,28 (0,69)	-0,41 (0,69)	0,03 (0,83)	1,00 (0,72)
2	$\theta = -2$	-2,18 (0,62)	<b><u>-1,75 (0,69)</u></b>	<b><u>-1,55 (0,70)</u></b>	<b><u>-1,96 (0,88)</u></b>	<b><u>-1,64 (1,18)</u></b>
	C	0,00 (0,01)	0,06 (0,07)	0,16 (0,10)	0,26 (0,12)	0,35 (0,13)
3	$\theta = -2$	-1,98 (0,64)	<b><u>-1,61 (0,74)</u></b>	<b><u>-1,46 (0,79)</u></b>	<b><u>-1,62 (1,10)</u></b>	<b><u>-1,22 (1,59)</u></b>
	C	0,01 (0,02)	<b><u>0,07 (0,08)</u></b>	<b><u>0,17 (0,11)</u></b>	<b><u>0,24 (0,14)</u></b>	<b><u>0,33 (0,17)</u></b>
	S	0,07 (0,23)	<b><u>0,17 (0,44)</u></b>	<b><u>0,23 (0,51)</u></b>	<b><u>0,44 (0,92)</u></b>	<b><u>0,43 (0,87)</u></b>
	D	0,03 (0,08)	0,03 (0,07)	0,03 (0,06)	0,03 (0,07)	0,02 (0,05)

<sup>1</sup>  $\theta = -2$ , C = 0,00 à 0.40, S = 0, D = 0, b = -5 à 5, c = 0, d = 0, 40 items, 100 sujets simulés  
 Modèle 1 :  $\theta$  uniquement; modèle 2 :  $\theta$  et C; modèle 3  $\theta$ , S, C et D





## APPLICATION À LA CORRECTION DU NIVEAU D'HABILITÉ (3)

**Tableau 2. Effet de l'utilisation des nouvelles modélisations sur la correction de l'estimation du niveau d'habileté de 100 sujets qui cherchent à augmenter leur résultat à un test de 40 items**

		<b>C (paramètre de pseudo chance personnelle)</b>				
<b>Modèle</b>		<b>0,00</b>	<b>0,10</b>	<b>0,20</b>	<b>0,30</b>	<b>0,40</b>
<b>1</b>	$\theta = 2$	2,19 (0,55)	2,52 (0,61)	3,02 (0,58)	3,31 (0,62)	3,46 (0,56)
<b>2</b>	$\theta = 2$	2,09 (0,61)	2,37 (0,60)	2,52 (0,85)	2,74 (0,88)	2.51 (1,01)
	C	0,02 (0,06)	0,04 (0,10)	0,12 (0,18)	<b>0,16 (0,21)</b>	<b>0,27 (0,23)</b>
<b>3</b>	$\theta = 2$	1,95 (0,55)	<b><u>2,11 (0,61)</u></b>	<b><u>2,10 (0,93)</u></b>	<b><u>2,42 (0,90)</u></b>	<b><u>2,02 (1,08)</u></b>
	S	0,11 (0,27)	0,17 (0,34)	0,12 (0,32)	0,12 (0,40)	0,09 (0,30)
	C	0,02 (0,05)	0,08 (0,13)	<b>0,20 (0,19)</b>	<b>0,22 (0,21)</b>	<b>0,35 (0,23)</b>
	D	0,01 (0,02)	0,00 (0,01)	0,00 (0,01)	0,00 (0,01)	0,00 (0,01)

<sup>1</sup>  $\theta = 2$ , C = 0,00 à 0.40, S = 0, D = 0, b = -5 à 5, c = 0, d = 0, 40 items, 100 sujets simulés

Modèle 1 :  $\theta$  uniquement; modèle 2 :  $\theta$  et C; modèle 3  $\theta$ , S, C et D



# APPLICATION À LA CORRECTION DU NIVEAU D'HABILITÉ (4)

**Tableau 3. Effet de l'utilisation des nouvelles modélisations sur la correction de l'estimation du niveau d'habileté de 100 sujets dont le niveau d'habileté varie anormalement à un test de 40 items**

		S (paramètre de variabilité personnelle)				
Modèle		0,00	0,50	1,00	2,00	4,00
1	$\theta = -2$	<b><u>-2,14 (0,59)</u></b>	<b><u>-2,07 (0,61)</u></b>	-2,00 (0,62)	-1,71 (0,70)	-1,01 (0,86)
2	$\theta = -2$	-2,06 (0,58)	-2,03 (0,59)	<b><u>-2,05 (0,67)</u></b>	<b><u>-2,00 (0,84)</u></b>	<b><u>-1,63 (1,28)</u></b>
	S	0,33 (0,42)	0,51 (0,57)	<b><u>0,90 (0,68)</u></b>	<b><u>1,87 (0,90)</u></b>	<b><u>3,45 (0,75)</u></b>
3	$\theta = -2$	-1,92 (0,66)	-1,75 (0,79)	-1,54 (0,83)	-0,97 (1,42)	-0,68 (2,37)
	S	0,11 (0,24)	0,15 (0,29)	<b><u>0,28 (0,45)</u></b>	<b><u>0,71 (0,84)</u></b>	<b><u>0,79 (1,24)</u></b>
	C	0,01 (0,01)	0,01 (0,02)	0,01 (0,03)	0,03 (0,05)	<b><u>0,15 (0,14)</u></b>
	D	0,04 (0,09)	0,06 (0,11)	<b><u>0,11 (0,14)</u></b>	<b><u>0,18 (0,18)</u></b>	<b><u>0,23 (0,21)</u></b>

<sup>1</sup>  $\theta = -2$ , C = 0,00 à 4,00, S = 0, D = 0, b = -5 à 5, c = 0, d = 0, 40 items, 100 sujets simulés  
 Modèle 1 :  $\theta$  uniquement; modèle 2 :  $\theta$  et C; modèle 3  $\theta$ , S, C et D



## APPLICATION À LA CORRECTION DU NIVEAU D'HABILITÉ (5)

**Tableau 4. Effet de l'utilisation des nouvelles modélisations sur la correction de l'estimation du niveau d'habileté de 100 sujets dont le niveau d'habileté varie anormalement à un test de 40 items**

		<b>S (paramètre de variabilité personnelle)</b>				
<b>Modèle</b>		<b>0,00</b>	<b>0,50</b>	<b>1,00</b>	<b>2,00</b>	<b>4,00</b>
<b>1</b>	$\theta = 2$	<b><u>2,01 (0,52)</u></b>	<b><u>1,97 (0,55)</u></b>	2,03 (0,66)	1,70 (0,73)	1,13 (0,84)
<b>2</b>	$\theta = 2$	1,98 (0,49)	1,92 (0,54)	<b><u>2,08 (0,71)</u></b>	<b><u>2,02 (0,90)</u></b>	<b><u>1,63 (1,16)</u></b>
	S	0,27 (0,46)	0,40 (0,49)	<b><u>0,97 (0,66)</u></b>	<b><u>1,93 (0,96)</u></b>	<b><u>3,19 (0,88)</u></b>
<b>3</b>	$\theta = 2$	1,88 (0,54)	1,77 (0,57)	1,62 (0,92)	1,20 (1,46)	0,38 (2,13)
	S	0,09 (0,29)	0,16 (0,36)	<b><u>0,27 (0,49)</u></b>	<b><u>0,53 (0,77)</u></b>	<b><u>0,38 (0,82)</u></b>
	C	0,03 (0,07)	0,04 (0,09)	<b><u>0,11 (0,14)</u></b>	<b><u>0,16 (0,18)</u></b>	<b><u>0,28 (0,18)</u></b>
	D	0,01 (0,02)	0,01 (0,02)	0,02 (0,04)	0,05 (0,07)	<b><u>0,13 (0,13)</u></b>

<sup>1</sup>  $\theta = 2$ , C = 0,00 à 4,00, S = 0, D = 0, b = -5 à 5, c = 0, d = 0, 40 items, 100 sujets simulés  
 Modèle 1 :  $\theta$  uniquement; modèle 2 :  $\theta$  et C; modèle 3  $\theta$ , S, C et D



# APPLICATION À LA CORRECTION DU NIVEAU D'HABILITÉ (6)

## Conclusions

- Les nouvelles modélisations n'apportent pas de biais, ni d'augmentation de l'erreur type, dans l'estimation du niveau d'habileté lorsque le patron de réponses est approprié
- Les nouvelles modélisations sont toutes assez efficaces pour corriger le niveau d'habileté
- Lorsque le niveau d'habileté est très faible et que la pseudo chance personnelle est présente, une modélisation comprenant uniquement la pseudo chance est la plus appropriée.
- Toutefois, lorsque le niveau d'habileté est élevé, la modélisation complète est la plus appropriée.
- Il est conseillé, du moins pour le moment, d'appliquer en premier lieu la modélisation complète. Ensuite, si le niveau d'habileté estimé est élevée, de conserver cette estimation. Toutefois, si le niveau d'habileté est faible, il est plutôt conseillé de conserver par la suite l'estimateur obtenu à partir du modèle avec pseudo chance
- Il faudra toutefois beaucoup plus de discernement dans l'utilisation de ces modélisations



# APPLICATION À LA CORRECTION DU NIVEAU D'HABILITÉ (7)

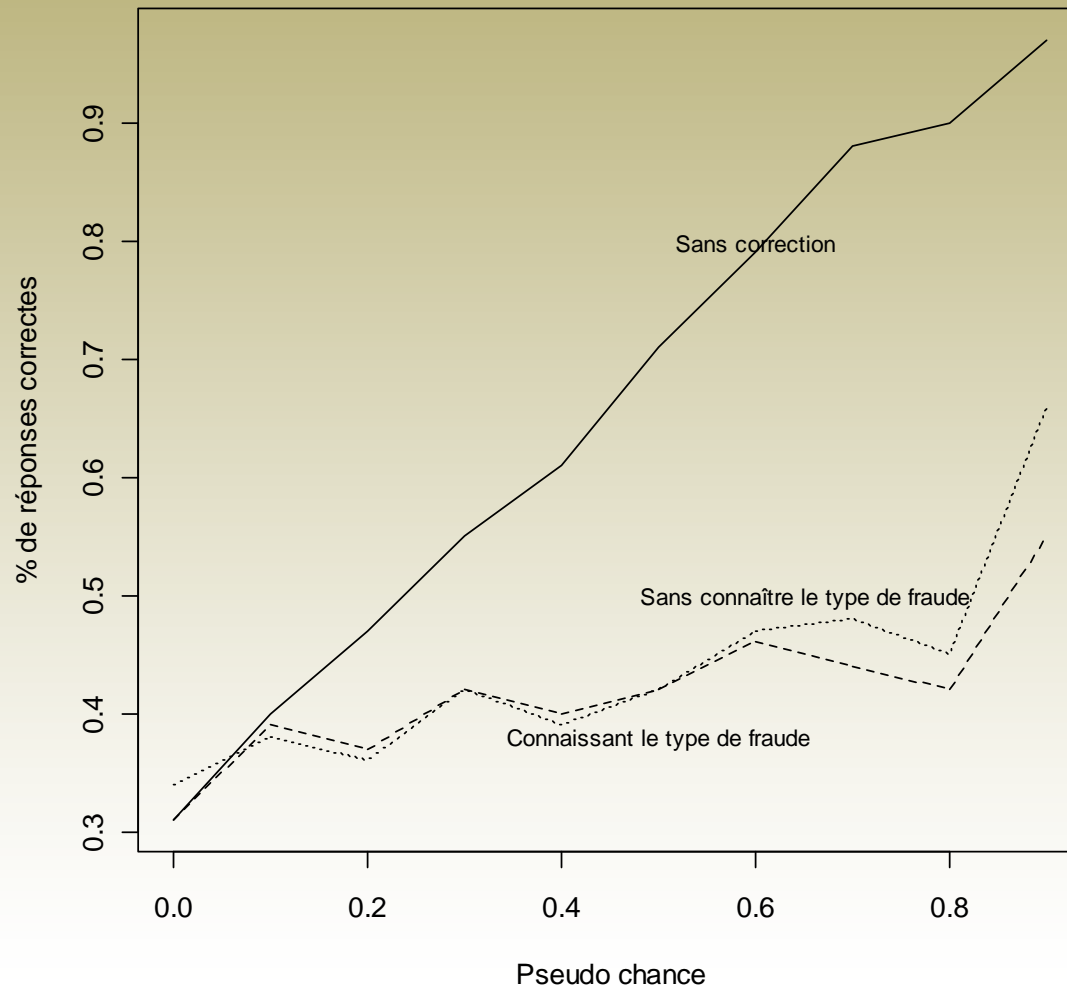
**Tableau 5. Comparaison de l'estimation de la proportion de réponses correctes selon trois modèles et dix valeurs du paramètre de pseudo chance personnelle**

Pseudo chance	Modèle sans tenir compte de la fraude		Modèle correct en tenant compte de la pseudo chance		Modèle choisi sans connaître préalablement le type de fraude	
	%	et	%	et	%	et
0.0	31	7	31	7	34	7
0.1	40	7	39	7	38	7
0.2	47	7	37	8	36	7
0.3	55	8	42	10	42	15
0.4	61	8	40	11	39	12
0.5	71	8	42	13	42	16
0.6	79	7	46	17	47	21
0.7	88	6	44	22	48	23
0.8	<b><u>90</u></b>	8	<b><u>42</u></b>	17	<b><u>45</u></b>	23
0.9	97	2	55	31	66	25



## APPLICATION À LA CORRECTION DU NIVEAU D'HABILITÉ (8)

Comparaison de l'estimation de la proportion de réponses correctes selon trois modèles et dix valeurs du paramètre de pseudo chance personnelle





# CONSIDÉRATIONS DANS L'UTILISATION DE CES MODÉLISATIONS

- **Leur comportement est encore peu connu**
- **Il est encore difficile de juger quel est le meilleur modèle à appliquer dans une situation spécifique**
- **Quel est l'impact de l'utilisation du mauvais modèle?**
- **Il faut augmenter l'étendue du niveau de difficulté des items au détriment de la précision du niveau d'habileté**
- **Il faut augmenter considérablement le nombre d'items si on désire une mesure précise des nouveaux paramètres de personne**
- **Il est préférable de planifier à l'avance un plan de mesure**
- **Il est préférable de miser sur des items à niveau élevé de discrimination**



## AUTRES APPLICATIONS

- **Simulation de patrons de réponses inappropriés**
- **Indices de détection de patrons de réponses inappropriés**
- **Étude de comportements spécifiques à des situations de test en lien avec ces modélisations**





## APPLICATION LOGICIELLE: *irtProb*

- **Implantée dans le logiciel *R***
- **Disponible directement sur le site de CRAN**
- **Fonctionnalités:**
  1. **Conversions des paramètres d'items entre TRI et TCT**
  2. **Fonctions de probabilité, de densité, de quantile et de simulation pour les modélisations de 1 à 4 paramètres**
  3. **Fonctions de probabilité, de densité, de quantile et de simulation pour les modélisations multidimensionnelles de 1 à 4 paramètres**
  4. **Fonctions d'estimation des paramètres de personne par MLE et MAP**
  5. **Représentations graphiques des courbes de vraisemblance**
  6. **Diverses fonctions utilitaires**



# RÉFÉRENCES

- Berg, I. A. (1955). Response bias and personality. *Journal of psychology*, 68(1), 60-71.
- Cronbach, L. J. (1950). Further evidence on response sets and test design. *Educational and psychological measurement*, 10(1), 3-31.
- Ferrando, P. J. (2004). Person reliability in personality measurement : an item response theory analysis. *Applied psychological measurement*, 28(2), 126-140.
- Fiske, D. W. and Rice, L. (1955). Intra-individual response variability. *Psychological bulletin*, 52(3), 217-250.
- Frederiksen, N. and Messick, S. (1959). Response set as a measure of personality. *Educational and psychological measurement*, 19(2), 137-157.
- Levine, M. V. and Drasgow, F. (1983). Appropriateness measurement : validating studies and variable ability models. Dans D. J. Weiss (Dir.) : *New horizons in testing. Latent trait test theory and computerized adaptive testing*. New York, New Jersey : Academic Press.
- Raïche, G., Blais, J.-G. et Magis, D. (2009). La détection des patrons de réponses problématiques dans le contexte des tests par ordinateur. Dans J.-G. Blais( Dir.) : *Évaluation des apprentissages et technologies de l'information et de la communication : enjeux, applications et modèles de mesure*. Sainte-Foy, Québec : Presses de l'Université Laval.
- Trabin, T. E. and Weiss, D. J. (1983). The person response curve : fit of individuals to item response theory models. Dans D. J. Weiss (Dir.) : *New horizons in testing. Latent trait test theory and computerized adaptive testing*. New York, New Jersey : Academic Press.
- Wiggins, J. S. (1962). Definition of social desirability and acquiescence in personality inventories. Dans S. M. Messick and J. Ross (Dir.) : *Measurement in personality and cognition*. New York, New Jersey : John Wiley.