

REVUE DE STATISTIQUE APPLIQUÉE

J.-M. TRICOT

M. MESBAH

Un modèle de réponses aux items. Propriétés et comparaison de groupes de traitement en épidémiologie

Revue de statistique appliquée, tome 48, n° 4 (2000), p. 29-39

http://www.numdam.org/item?id=RSA_2000__48_4_29_0

© Société française de statistique, 2000, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

UN MODÈLE DE RÉPONSES AUX ITEMS. PROPRIÉTÉS ET COMPARAISON DE GROUPES DE TRAITEMENT EN ÉPIDÉMIOLOGIE

J.-M. Tricot*, M. Mesbah*

* *Laboratoire de Statistique Appliquée Bretagne-Sud*

Université de Bretagne-Sud, IUP Informatique & Statistique, Tohannic, 56000 Vannes

RÉSUMÉ

La Théorie des Réponses aux Items (IRT) est étudiée dans le cadre des items à réponses dichotomiques. Le modèle de Rasch possède de bonnes propriétés parmi les modèles IRT unidimensionnels. On s'intéresse particulièrement au cas où on a un modèle à effet aléatoire. On présente la combinaison du modèle de Rasch et d'un modèle linéaire sur la variable latente (effet aléatoire) avec application au cas où les variables prédictrices sont des variables d'affectation. Dans ce contexte, un exemple réel porte sur la comparaison de groupes de traitements en diabétologie. Le test de la comparaison des groupes est effectué à l'aide de la statistique LR.

Mots-clés : *Modèles IRT, Modèles de Rasch, Trait latent, Exhaustivité, Score, Variables prédictrices, Statistique LR*

ABSTRACT

The Item Response Theory (IRT) is analysed Within the framework of dichotomous responses items. Rasch model have some good properties that make them outstanding univariate IRT model. The random effect model is of our interest. We insert a linear model applied to the latent variable (random effect) inside the Rasch model with binary predictors to classes' assignment. Within this framework, we analyse two treatments groups comparison as an example in a diabetes study. The group's comparison test is a LR test.

Keywords : *IRT models, Rasch model, Latent trait, Sufficiency, Score, Predictors, LR Statistics*

1. Introduction

On est souvent amené à analyser des enquêtes où les données ont été obtenues par autoquestionnaire. Les autoquestionnaires sont couramment utilisés dans la recherche médicale pour mesurer des variables telles que la douleur, l'émotion etc...

Plus généralement l'analyse des autoquestionnaires est réalisée dans le domaine de la Qualité de Vie (Pfanzagl, 1994).

La Théorie des Réponses aux Items (IRT) est développée dans ce cadre. Une revue rapide des modèles IRT est donnée par Andersen (1991). Chaque autoquestionnaire est assimilable à un ensemble d'items dont on va supposer qu'ils sont dichotomiques et qu'ils mesurent un unique critère d'évaluation ou trait latent : ceci revient à dire que la variable (critère ou trait latent) à mesurer est unidimensionnelle.

On se focalise sur un modèle dérivé du modèle de Rasch (1960) pour lequel l'ensemble des items représentent la partie instrumentale (mesurable) du modèle dont les paramètres sont des paramètres d'item. L'effet sujet, c'est-à-dire le trait latent (critère à mesurer) est considéré soit sous une forme paramétrique soit sous la forme d'un effet aléatoire : il représente la partie structurelle du modèle. Dans la partie 2, on définit le modèle et on fait un lien avec les méthodes classiques d'analyse de variance. Dans la partie 3, les propriétés essentielles du modèle sont abordées. La partie 4 présente un cadre théorique d'estimation des paramètres du modèle. La partie 5 traite de la combinaison du modèle de Rasch et d'un modèle linéaire que l'on restreint au cas où les prédicteurs sont des variables d'affectation. En 6^e partie, un exemple est donné dans le domaine de l'épidémiologie.

2. Notations et Modèles

L'utilisation du modèle de Rasch (1960) dichotomique est pertinente en présence d'un trait latent dans la population de sujets : on estime et teste une variable non observable ou variable latente («latent variable») à l'aide d'un ensemble de variables observables ou items («manifest variables»). L'ensemble des items constitue l'autoquestionnaire. C'est un instrument de mesure pour la variable latente. Les items sont, ici, des variables qualitatives dichotomiques (réponse 0 ou 1). Dans le domaine de la psychométrie par exemple (cf. la 6^e partie), l'état émotionnel d'un sujet est mesuré sur des réponses à des questions objectives du type «je me sens nerveux, tendu», «la vie ne vaut plus la peine d'être vécue» etc... Ces variables dichotomiques admettent les réponses «oui» (échec), ou «non» (succès). Les codages respectifs sont 0 et 1, par exemple. Soient n sujets et k questions (items). La question X_i (ou plus simplement i), $i = 1, k$, admet la réponse x_i , $x_i = 0, 1$. Ces valeurs sont considérées comme un système de codage sur les réponses.

On distingue deux parties dans le modèle. D'une part, une partie instrumentale qui concerne les items et qui est caractérisée par un ensemble de paramètres d'item : le paramètre de l'item i , est noté β_i , $i = 1, k$. Il est interprété comme étant la difficulté de l'item i : il s'agit de la difficulté à réussir i (dans l'exemple, si l'item i est : «je me sens nerveux, tendu», la difficulté β_i mesure la difficulté à répondre $x_i = 1$). On suppose que $\beta_i \in \mathbf{R}$. D'autre part, une partie structurelle qui concerne le trait latent (ou variable latente) noté θ . Il est interprété comme étant l'aptitude du sujet à réussir l'item i , $i = 1, k$, c'est-à-dire à répondre $x_i = 1$. Si l'on spécifie le sujet à l'aide de l'indice s , $s = 1, n$, on passe de manière évidente de la notation θ à θ_s (aptitude du sujet s) et de X_i et x_i respectivement à X_{si} et x_{si} . Suivant le point de vue adopté, on supposera que $\theta \in \mathbf{R}$ est un paramètre fixe ou un paramètre aléatoire (une variable aléatoire réelle).

Classiquement, la modélisation d'un effet item (effet fixe donné par β_i) et d'un effet sujet (effet fixe ou aléatoire donné par θ_s) sur la réalisation x_{si} est obtenu linéairement à l'aide d'un plan à deux facteurs sans répétition (ou à un facteur «avec blocs». Voir Tomassone *et al.*, 1993), donné par :

$$X_{si} = \theta_s + \beta_i + \varepsilon_{si}, \quad (1)$$

où θ_s est un effet sujet et β_i est un effet item.

Plusieurs travaux récents sur la Théorie des Réponses aux Items se focalisent non plus sur la variable réponse X_{si} mais plutôt sur une transformation de la probabilité $P(X_{si} = 1 | \theta_s, \beta_i)$ dans le cas dichotomique. Dans ce cadre, le modèle linéaire $P(X_{si} = 1 | \theta_s, \beta_i) = \theta_s + \beta_i$ n'est pas satisfaisant puisque des valeurs supérieures à 1 ou négatives pourraient être prises par le membre de droite, à moins d'imposer à celui-ci une contrainte.

Le modèle logistique de Rasch retenu récemment par de nombreux auteurs, intègre une telle contrainte : il est construit à partir de la transformation logit appliquée à la probabilité $P(X_{si} = 1 | \theta_s, \beta_i)$:

$$\text{logit}P(X_{si} = 1 | \theta_s, \beta_i) = \theta_s - \beta_i, \quad (2)$$

où le signe négatif devant β_i permet d'opposer l'aptitude du sujet (θ_s) à la difficulté de l'item (β_i). D'autres liens ont aussi été utilisés en pratique. Entre autre on peut citer le lien probit (par exemple, cf. McCullagh & Nelder, 1989).

La fonction IRF («Item Response Function») de l'item i , est la fonction qui, pour ce modèle logistique, à θ associe $P(X_i = 1 | \theta, \beta_i)$, $i = 1, k$. Les k courbes logistiques associées aux k fonctions IRF, sont parallèles puisqu'elles ne dépendent des effets sujet et item que *via* la différence $\theta - \beta_i$, $i = 1, k$. L'équation (2) et les notations introduites permettent d'écrire :

$$P(X_{si} = x_{si} | \theta_s, \beta_i) = \frac{\exp(x_{si}(\theta_s - \beta_i))}{1 + \exp(\theta_s - \beta_i)}, \quad (3)$$

où x_{si} prend la valeur 0 ou 1.

3. Définitions et propriétés

Le modèle de Rasch dichotomique est défini par les deux caractéristiques suivantes :

- 1) la forme de la fonction IRF donnée par l'équation (2).
- 2) l'indépendance locale des items, c'est-à-dire : pour un individu donné (θ fixé), les X_i sont des variables indépendantes (pas d'empiètement des questions) et de plus on a l'unidimensionalité de θ ($\theta \in \mathbf{R}$) et par conséquent l'aptitude θ suffit à expliquer toutes les réponses.

Le modèle logistique de Rasch appartient à la famille exponentielle en θ et β . On en déduit donc immédiatement, lorsque θ_s est considéré comme un paramètre fixe,

l'exhaustivité des statistiques $R_s = \sum_i X_{si}$ pour θ_s (score pour θ_s), et $T_i = \sum_s X_{si}$ pour β_i (score pour β_i). On prend naturellement la notation R pour désigner le score pour θ . Fischer (1995) montre que, sous de bonnes conditions de régularité, il est équivalent, pour un modèle IRT unidimensionnel, de supposer que c'est un modèle de Rasch ou de supposer les trois hypothèses suivantes : la monotonie croissante des courbes des fonctions IRT, l'indépendance locale des items et l'exhaustivité du score R pour θ .

Pour un tel modèle logistique, il existe un «comparateur» entre deux sujets s et s' , indépendant du modèle de mesure (indépendant des β), noté $C(s, s')$, défini, quelque soit i , par :

$$C(s, s') = P(X_{si} = 1 \wedge X_{s'i} = 0 | X_{si} + X_{s'i} = 1). \quad (4)$$

La définition (4) donne, par un calcul élémentaire, la valeur de $C(s, s')$, indépendante des β , suivante :

$$C(s, s') = 1/(1 + \exp(\theta_{s'} - \theta_s)). \quad (5)$$

En particulier, d'après (5), $C(s, s') \geq 1/2$ si et seulement si, quelque soit i , $\theta_{s'} \leq \theta_s$ ou si et seulement si $P(X_{s'i} = 1) \leq P(X_{si} = 1)$. On voit immédiatement qu'une condition suffisante pour que $\theta_{s'} \leq \theta_s$ est qu'il existe i tel que $P(X_{s'i} = 1) \leq 1/2$ (i.e. $\theta_{s'} \leq \beta_i$) et $P(X_{si} = 1) \geq 1/2$ (i.e. $\theta_s \geq \beta_i$). Un item suffit donc à classer 2 sujets. En ce sens, chaque β , représenté sur l'échelle unidimensionnelle des valeurs réelles de θ , est valeur discriminante pour les sujets dans la détermination du classement de l'ensemble des sujets au sens d'une aptitude «de plus en plus grande» à réussir les items.

4. Estimation des paramètres

On note X_s , le vecteur des réponses du sujet s aux items, à k composantes, de $i^{\text{ème}}$ composante X_{si} , $i = 1, k$, et \mathbf{x}_s une réalisation de \mathbf{X}_s , de $i^{\text{ème}}$ composante x_{si} , $i = 1, k$. On note aussi $\beta = (\beta_1, \dots, \beta_k)$. L'exhaustivité du score R_s (cas où θ_s est considéré comme un paramètre fixe) permet d'effectuer la factorisation suivante, pour une réalisation r_s de R_s :

$$P(\mathbf{X}_s = \mathbf{x}_s | \theta_s, \beta) = P(R_s = r_s | \theta_s, \beta) P(\mathbf{X}_s = \mathbf{x}_s | r_s, \beta), \quad (6)$$

où le deuxième facteur de la partie droite de l'équation ne dépend pas de θ_s .

On a : $P(R_s = r_s | \theta_s, \beta) = \sum_{\mathbf{x}_s} P(R_s = r_s | \mathbf{x}_s, \theta_s, \beta) P(\mathbf{X}_s = \mathbf{x}_s | \theta_s, \beta)$ où la somme est réalisée sur 2^k valeurs possibles de \mathbf{x}_s , pour s fixé. Du fait que, de manière évidente, on a $P(R_s = r_s | \mathbf{x}_s, \theta_s, \beta) = 1$ si \mathbf{x}_s est tel que $r_s = \sum_i x_{si}$ (condition

notée $\mathbf{x}_s|r_s$) et 0 sinon, la loi de R_s est donc donnée par :

$$P(R_s = r_s|\theta_s, \beta) = \sum_{\mathbf{x}_s|r_s} \frac{\exp(r_s\theta_s - \sum_i x_{si}\beta_i)}{\prod_i (1 + \exp(\theta_s - \beta_i))}, \quad (7)$$

et, par conséquent, la loi conditionnelle de \mathbf{X}_s sachant r_s est donnée par

$$P(\mathbf{X}_s = \mathbf{x}_s|r_s, \beta) = \frac{\exp(-\sum_i x_{si}\beta_i)}{\sum_{\mathbf{x}_s|r_s} \exp(-\sum_i x_{si}\beta_i)}, \quad (8)$$

si r_s est tel que $r_s = \sum_i x_{si}$ et 0 sinon.

On se place maintenant dans le cas d'un modèle mixte où θ est aléatoire, de loi $F(\cdot|\lambda)$ paramétrée par λ (λ sera, par exemple le paramètre (μ, σ^2) d'une loi normale).

Plusieurs auteurs et en particulier Zwinderman (1991) donnent les étapes d'estimation suivantes :

Première étape : estimation du paramètre β par la méthode CML, c'est-à-dire en maximisant la log-vraisemblance de β , notée L_C , calculée à partir de (8) :

$$L_C = \log \prod_s P(\mathbf{X}_s = \mathbf{x}_s|r_s, \beta) = \sum_s \log \left[\frac{\exp(-\sum_i x_{si}\beta_i)}{\sum_{\mathbf{x}_s|r_s} \exp(-\sum_i x_{si}\beta_i)} \right]. \quad (9)$$

L'estimation de β est obtenue pratiquement par les méthodes algorithmiques usuelles (procédures EM ou Newton-Raphson). On consultera, par exemple, Fischer et Ponocny (1994) pour une description de ces méthodes.

Deuxième étape : estimation du paramètre λ par la méthode MML, c'est-à-dire en maximisant la log-vraisemblance de λ , notée L_M , lorsque β est supposé connu, c'est-à-dire en remplaçant β (resp. β_i , $i = 1, k$) par son estimation $\hat{\beta}$ (resp. $\hat{\beta}_i$, $i = 1, k$) obtenue à la première étape (cf. (9)). Pour cela, on écrit L_M à l'aide des équations (6) et (7) :

$$L_M = \log \prod_s \int P(\mathbf{X}_s = \mathbf{x}_s|\theta, \hat{\beta}) dF(\theta|\lambda) = \sum_s \log P(\mathbf{X}_s = \mathbf{x}_s|r_s, \hat{\beta}) + \sum_s \log \int \sum_{\mathbf{x}_s|r_s} \frac{\exp(r_s\theta - \sum_i x_{si}\hat{\beta}_i)}{\prod_i (1 + \exp(\theta - \hat{\beta}_i))} dF(\theta|\lambda). \quad (10)$$

Seul le deuxième terme de la somme (10) précédente dépend de λ . C'est ce terme qui est maximisé et permet d'obtenir une estimation $\hat{\lambda}$ de λ .

L'estimation $\hat{\theta}_s$ de θ_s est alors l'espérance mathématique *a posteriori* de θ sachant $\mathbf{X}_s = \mathbf{x}_s$, lorsque β et λ sont estimés respectivement par $\hat{\beta}$ et $\hat{\lambda}$. La loi *a posteriori* de θ dans le cas absolument continu où $f(\theta|\lambda)$ est la densité *a priori* de θ , est donnée en fonction de $P(\mathbf{X}_s = \mathbf{x}_s|\theta, \beta)$ (cf. (6)) et $f(\theta|\lambda)$, par la densité, notée $f_s(\theta|\beta, \lambda)$ suivante (formule de Bayes habituelle) :

$$f_s(\theta|\beta, \lambda) = B^{-1}(\beta, \lambda) \frac{\exp\left(-\sum_i x_{si}\hat{\beta}_i\right)}{\prod_i (1 + \exp(\theta - \hat{\beta}_i))} \exp(r_s\theta) f(\theta|\lambda), \quad (11)$$

$$\text{où } r_s = \sum_i x_{si} \text{ et } B(\beta, \lambda) = \int \frac{\exp\left(-\sum_i x_{si}\hat{\beta}_i\right)}{\prod_i (1 + \exp(\theta - \hat{\beta}_i))} \exp(r_s\theta) f(\theta|\lambda) d\theta.$$

De (11) on déduit l'estimation de θ_s : $\hat{\theta}_s = \int \theta f_s(\theta|\hat{\beta}, \hat{\lambda}) d\theta$. Dans le cadre du modèle de Rasch à effet mixte, on constate donc qu'il est possible d'obtenir un intervalle de confiance autour de θ_s puisque l'on peut calculer la variance conditionnelle *a posteriori* de θ connaissant le score individuel du sujet s .

5. Combinaison d'un modèle de Rasch et d'un modèle linéaire sur θ

On se place dans le cas où la variable latente θ est aléatoire et supposée liée linéairement à P prédicteurs z_1, z_2, \dots, z_P . Pour le sujet s , les observations sur les P prédicteurs sont $z_{s1}, z_{s2}, \dots, z_{sP}$. On pose donc :

$$\theta_s = \gamma_0 + \gamma_1 z_{s1} + \dots + \gamma_P z_{sP} + \varepsilon_s, \quad (12)$$

en supposant que les ε_s , $s = 1, n$, sont des erreurs normales $N(0, \sigma^2)$ homoscédastiques.

Soit $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_P)$. On considère le modèle qui combine un modèle de Rasch et le modèle linéaire sur θ défini en (12), suivant :

$$P(X_{si} = x_{si} | \varepsilon_s, \gamma, \beta_i) = \frac{\exp(x_{si}(\gamma_0 + \gamma_1 z_{s1} + \dots + \gamma_P z_{sP} + \varepsilon_s - \beta_i))}{1 + \exp(\gamma_0 + \gamma_1 z_{s1} + \dots + \gamma_P z_{sP} + \varepsilon_s - \beta_i)}. \quad (13)$$

Il s'agit d'un modèle de Rasch à effet mixte (Hoijtink, 1995) qui combine un modèle de mesure à effet fixe (les β sont les paramètres du questionnaire qui sert d'instrument de mesure) et un modèle structurel à effet aléatoire linéaire (θ est la variable qui décrit la structure de la population de sujet).

On particularise le problème au cas où les prédicteurs sont des variables indicatrices d'affectation à des groupes de traitement. Fischer et Seliger (1997) étudient une telle comparaison de groupes de traitement effectuée lorsque les groupes sont observés à différentes périodes du temps.

Une étude épidémiologique récente, EDIAB (éducation des diabétiques : on pourra consulter par exemple, Chwalow *et al.*, 1994), considérait deux groupes de traitement, et avait pour but le «diagnostic et l'évaluation de stratégies éducatives pour les diabétiques non insulino-dépendants (DNID)».

Soit z_g , $g = 1, 2$, la variable indicatrice d'affectation au groupe de traitement g .

On suppose que, pour le groupe 1, l'indice sujet varie comme suit : $s = 1, n_1$, et que dans ce groupe l'aptitude moyenne des patients est μ_1 . Dans le groupe 2, $s = n_1 + 1, n_1 + n_2$ et l'aptitude moyenne des patients est μ_2 .

L'aptitude du patient s est alors $\theta_s = \mu_1 + \varepsilon_s$ lorsque $s = 1, n_1$ où l'on suppose que ε_s possède une distribution $N(0, \sigma_1^2)$ et de même $\theta_s = \mu_2 + \varepsilon_s$ lorsque $s = n_1 + 1, n_1 + n_2$ où l'on suppose que ε_s possède une distribution $N(0, \sigma_2^2)$. On admet que le paramètre de difficulté pour l'item i est β_i^1 (resp. β_i^2) dans le premier groupe (resp. le deuxième groupe) et que le paramètre de difficulté pour cet item tout groupe confondu (donc, pour les $n_1 + n_2$ sujets), est β_i^0 (plutôt que β_i). L'équation (13) redéfinie lorsque $P = 1$ et $\gamma_0 = 0$ dans le modèle (12), est équivalente, pour le groupe 1, à :

$$P(X_{si} = x_{si} | \varepsilon_s, \mu_1, \beta_i^1) = \frac{\exp(x_{si}(\mu_1 + \varepsilon_s - \beta_i^1))}{1 + \exp(\mu_1 + \varepsilon_s - \beta_i^1)}, s = 1, n_1, \quad (14)$$

et pour le groupe 2, à

$$P(X_{si} = x_{si} | \varepsilon_s, \mu_2, \beta_i^2) = \frac{\exp(x_{si}(\mu_2 + \varepsilon_s - \beta_i^2))}{1 + \exp(\mu_2 + \varepsilon_s - \beta_i^2)}, s = n_1 + 1, n_1 + n_2. \quad (15)$$

Dans ce contexte, on suppose que la distribution $F(\cdot | \lambda)$ de θ , dans le groupe 1, est normale $N(\mu_1, \sigma_1^2)$ avec $\lambda = (\mu_1, \sigma_1^2)$ et celle de θ , dans le groupe 2, est normale $N(\mu_2, \sigma_2^2)$ avec $\lambda = (\mu_2, \sigma_2^2)$.

6. Exemple et test de l'identité de deux groupes de traitement

Une enquête auprès des médecins et leurs malades a permis de collecter 1 366 réponses de patients à un questionnaire mesurant outre des informations d'ordre socio-démographiques (age, sexe, CSP, etc...) des items de santé (niveau de contrôle glycémique etc...), des échelles cognitives, psycho-sociales et enfin une échelle multidimensionnelle de qualité de vie : l'échelle Nottingham Health Profile (NHP) constituée par 38 items dichotomiques regroupant 6 dimensions ou sous-échelles : la douleur (8 items), les réactions émotionnelles (9 items), l'énergie/le tonus (3 items), la mobilité physique (8 items), l'isolement social (5 items) et le sommeil (5 items). On s'intéresse à comparer les réactions émotionnelles dans deux groupes de patients, suivant le niveau de contrôle glycémique. Le niveau de contrôle glycémique est élevé dans le groupe 1 et faible dans le groupe 2.

Les 9 items dichotomiques de l'échelle des réactions émotionnelles sont les suivants :

- Item 1 : Je suis de plus en plus découragé
- Item 2 : Plus rien ne me fait plaisir
- Item 3 : Je me sens nerveux, tendu
- Item 4 : Je trouve les journées interminables
- Item 5 : Je me mets facilement en colère
- Item 6 : J'ai du mal à faire face aux événements
- Item 7 : Mes soucis m'empêchent de dormir
- Item 8 : La vie ne vaut pas la peine d'être vécue
- Item 9 : Je me réveille déprimé le matin

La réponse à un item est codée 1 si celle-ci est négative (dans ce cas la réaction émotionnelle est contrôlée et la réponse «non» est un succès), sinon elle est codée 0.

Compte tenu de la présence de données manquantes, on peut analyser 1 135 patients répartis en deux groupes de taille 574 pour le groupe 1 et 561 pour le groupe 2. On se place dans le cadre du modèle de Rasch à effet mixte. On note de manière évidente : $\beta^0 = (\beta_1^0, \dots, \beta_k^0)$ et $\beta^g = (\beta_1^g, \dots, \beta_k^g)$ pour $g = 1, 2$ et $k = 9$. Le test de l'identité des groupes réalisé, est le test *LR* du rapport de vraisemblance. On teste l'égalité des paramètres, soit $H_0 : \mu_1 = \mu_2 = \mu_0$ et $\sigma_1^2 = \sigma_2^2 = \sigma_0^2$ et $\beta^1 = \beta^2 = \beta^0$, contre l'alternative contraire.

La significativité du test pourra intervenir aussi bien dans le cas d'un effet traitement positif sur les groupes que dans le cas où l'instrument de mesure suffit à séparer les groupes.

L'estimation, par la méthode CML, des paramètres β^1 et β^2 respectivement dans chacun des groupes 1 et 2, donne les résultats reportés dans les tableaux 1 et 2 suivants :

TABLEAU 1
*Estimation de la difficulté des items dans le groupe 1
et score des items (0, réponse oui, 1, réponse non).*

Item	difficulté $\hat{\beta}_i^1$	écart-type	score
Item 1	0,322	0,15	508
Item 2	0,505	0,15	515
Item 3	- 1,808	0,12	382
Item 4	0,861	0,17	527
Item 5	- 1,186	0,12	426
Item 6	0,11	0,14	499
Item 7	- 0,455	0,13	471
Item 8	1,031	0,18	532
Item 9	0,616	0,13	519

TABLEAU 2
 Estimation de la difficulté des items dans le groupe 2
 et score des items (0, réponse oui, 1, réponse non).

Item	difficulté $\hat{\beta}_i^2$	écart-type	score
Item 1	0,071	0,13	466
Item 2	0,411	0,14	483
Item 3	- 1,578	0,12	354
Item 4	0,916	0,16	504
Item 5	- 1,124	0,12	389
Item 6	0,053	0,13	465
Item 7	- 0,657	0,12	422
Item 8	1,41	0,18	520
Item 9	0,499	0,17	487

On pose :

$$\psi_{sh}(\mu_h, \sigma_h^2, \beta^h) = \int P(\mathbf{X}_s = \mathbf{x}_s | \theta, \beta^h) \frac{1}{\sigma_h \sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_h^2}(\theta - \mu_h)^2\right) d\theta, \quad (16)$$

pour $h = 0, 1, 2$.

Les trois vraisemblances calculées pour l'ensemble des sujets et pour chacun des deux groupes sont notées respectivement L_0 , L_1 (pour le groupe 1) et L_2 (pour le groupe 2). Les équations (10), (14), (15) et (16) permettent alors d'écrire :

$$L_0 = \prod_{s=1}^{n_1+n_2} \psi_{s0}(\mu_0, \sigma_0^2, \beta^0), L_1 = \prod_{s=1}^{n_1} \psi_{s1}(\mu_1, \sigma_1^2, \beta^1)$$

$$\text{et } L_2 = \prod_{s=n_1+1}^{n_1+n_2} \psi_{s2}(\mu_2, \sigma_2^2, \beta^2).$$

On note \hat{L}_0 , \hat{L}_1 et \hat{L}_2 les maxima respectivement de L_0 , L_1 et L_2 aux points qui donnent les estimations des paramètres de position et d'échelle. La statistique LR du test de l'identité des deux groupes de traitement, à savoir :

$$LR = 2(\log \hat{L}_0 - \log \hat{L}_1 - \log \hat{L}_2), \quad (17)$$

possède asymptotiquement, lorsque $n \rightarrow \infty$, une distribution χ^2 à $k + 2$ degrés de liberté.

Le logiciel RSP de Glas (1993) donne les valeurs : $\log \hat{L}_0 = -3\,664.4$, $\log \hat{L}_1 = -1\,713.2$, $\log \hat{L}_2 = -1\,940.8$. On remarque que les vraisemblances \hat{L}_0 , \hat{L}_1 et \hat{L}_2 prennent des valeurs très faibles indiquant que l'ajustement du modèle de Rasch n'est pas satisfaisant dans les trois populations (global, groupe 1 et groupe 2).

Nous n'avons pas abordé ici le problème de l'ajustement préalable du modèle de Rasch. De (17) on déduit la valeur $LR = 20.8$ pour 11 dl ce qui donne une p -value de .015. Au seuil de 5% on admettra une différence entre les deux groupes. On a aussi les valeurs moyennes suivantes : $\hat{\mu}_1 = -3.013(2.17)$ dans le groupe 1 est inférieure à $\hat{\mu}_2 = -2.460(1.99)$ dans le groupe 2. Puisque le groupe 1 correspond à un niveau de contrôle glycémique plus élevé que dans le groupe 2, cela signifie qu'une moyenne faible traduit un niveau de contrôle glycémique élevé : l'échelle de mesure du trait latent indique donc en moyenne des réactions émotionnelles moins fortes dans le deuxième groupe. On remarque enfin un bon ajustement du modèle au sens où les items les plus «difficiles» (pour des β faibles) ont les scores les moins élevés dans chacun des groupes : il s'agit des items 3 et 5. Le score de l'item 3 est de 382 dans le groupe 1 (le score maximum est 574) et de 354 dans le groupe 2 (le score maximum est 561). Pour l'item 5, les scores sont respectivement de 426 dans le groupe 1 et de 389 dans le groupe 2. La nervosité et la colère sont donc des attitudes facilement présentes quelque soit le niveau de contrôle glycémique.

7. Conclusion

Le test de comparaison des groupes de traitement a donné des résultats significatifs. Par rapport à une analyse de variance classique, le modèle de Rasch à effet mixte a montré sa pertinence et comment il pouvait être considéré comme une généralisation d'un modèle linéaire. L'importance de supposer l'effet sujet comme aléatoire permet d'envisager une variabilité au niveau du comportement du sujet. Cependant ce contexte n'est pas exempt de difficultés d'estimation des paramètres du modèle au niveau de la résolution numérique de problèmes d'optimisation. Le modèle de Rasch est performant dans la mesure où les outils algorithmiques sont maîtrisés.

L'hypothèse alternative posée ici, ne spécifie pas un modèle de Rasch. Sous cette hypothèse, les paramètres β dépendent du groupe considéré et donc des sujets mais en revanche, un modèle de Rasch particulier est spécifié par groupe. L'instrument utilisé sous l'hypothèse alternative, dépend donc des individus ce qui contredit le modèle de Rasch. Si l'on voulait ne pas altérer l'instrument de mesure considéré sous H_0 on prendrait l'alternative suivante : $\mu_1 \neq \mu_2$ ou $\sigma_1^2 \neq \sigma_2^2$ et $\beta^1 = \beta^2 = \beta^0$. Ainsi on pourrait tester la différence des groupes en terme d'aptitude des sujets et rester dans le cadre du modèle de Rasch. Les logiciels actuellement existant ne permettent pas d'utiliser directement leurs résultats pour construire ce test.

L'analyse, par le modèle de Rasch à effet mixte, de problèmes liés à la Qualité de Vie (dans notre cas, une étude épidémiologique) pourrait être explicitée de manière plus approfondie en vue de fournir une interprétation de la position réciproque des paramètres de moyenne dans les populations de sujet et des paramètres de difficulté des items.

Enfin, il faut préciser qu'il serait intéressant d'étendre l'étude précédente à plus de deux groupes de traitement et à des items polychotomiques. Les modèles de Rasch utilisés dans ce dernier cas sont alors de type polychotomique.

Bibliographie

- ANDERSEN E.B. (1991), *The Statistical Analysis of Categorical Data*, Heidelberg : Springer Verlag.
- CHWALOW A.J., COSTAGLIOLA D.G., MESBAH M., ESCHWEGE E. (1994), Management of type 2 diabetes mellitus in France : attitudes and practices among a representative sample of general practitioners, *Diabete & Metabolisme (Paris)*, **20**, 458-464.
- FISCHER G.H. (1995), in Fischer G.H. & Molenaar I.W., *Derivations of the Rasch Model*, New York : Springer Verlag.
- FISCHER G.H., PONOCNY I. (1994), An extension of the Partial Credit Model with an application to the measurement of change, *Psychometrika*, **59**, 177-192.
- FISCHER G.H., SELIGER E.S. (1997), in Van der Linden W.J. & Hambleton R.K., *Handbook of Modern Item Response Theory*, New York : Springer Verlag.
- GLAS C.A.W. (1993), *RSP, Rasch Scaling Program : Computer Program and User's Manual*, Groningen : ProGAMMA.
- HOIJTINK H. (1995), in Fischer G.H. & Molenaar I.W., *Linear and Repeated Measures Models for Person Parameters*, New York : Springer Verlag.
- MCCULLAGH P., NELDER J.A. (1989), *Generalized Linear Models*, London, Chapman and Hall.
- PFANZAGL J. (1994), in Fischer G.H. & Laming D., *Contributions to Mathematical Psychology, Psychometrics and Methodology*, New York : Springer Verlag.
- RASCH G. (1960), *Probabilistic Models for Some Intelligence and Attainment Tests*, Chicago, The University of Chicago Press.
- TOMASSONE R., DERVIN C., MASSON J.P. (1993), *Biométrie. Modélisation de Phénomènes Biologiques*, Paris, Masson.
- ZWINDERMAN A.H. (1991), A generalized Rasch Model for manifest predictors, *Psychometrika*, **56**, 589-600.