Application d'un nouveau modèle basé sur trois hypothèses d'invariance d'échelle à des séquences d'identification de photographies d'oiseaux et de constellations

Riopel, Martin, riopel.martin@uqam.ca, Université du Québec à Montréal, Canada
Chastenay, Pierre, chastenay.pierre@uqam.ca, Université du Québec à Montréal, Canada
Roy, Laval, lavar@videotron.ca, desoiseauxsurmaroute.blogspot.com, Canada
Choquette, Louise, Université du Québec à Montréal, Canada

L'évaluation continue en contexte scolaire présente plusieurs avantages. On peut considérer, par exemple, la possibilité pour les enseignants de personnaliser le processus de formation ou d'accompagnement afin de favoriser les apprentissages. On peut aussi considérer, pour les apprenants, la possibilité d'obtenir des rétroactions plus fréquentes et mieux adaptées susceptibles de donner plus de sens aux activités proposées. Cependant, la mise en application d'une procédure d'évaluation continue pose aussi plusieurs défis. Parmi ceux-ci, on peut mentionner la construction d'un cadre conceptuel et fonctionnel cohérent pour modéliser la progression. Ce cadre doit aussi présenter un niveau de précision suffisant pour guider l'optimisation à long terme des apprentissages. Les modèles habituellement considérés reposent sur la théorie des réponses aux items et présentent des difficultés techniques qui rendent difficile leur utilisation systématique par les praticiens. Afin de favoriser cette utilisation, il est nécessaire de poursuivre le développement de ces modèles pour les rendre mieux adaptés au contexte particulier de l'éducation. Ce contexte de développement peut être associé à l'édumétrie qui, selon la proposition de de Landsheere (1988, p. 59), correspond à l' « étude quantitative des variables relatives aux apprentissages suscités par l'éducation ».

Dans ce contexte de l'édumétrie, le modèle choisi doit permettre de suivre explicitement les apprentissages de façon aussi générale que possible. Ces cas correspondent à la famille des modèles à réponses ordonnées. Ces réponses pourront être évaluées selon une échelle dichotomique (par exemple : 0 ou 1), graduelle ou partielle (par exemple : 0/3, 1/3, 2/3, 3/3) ou séquentielle (par exemple : 0/1, 1/2, 2/3, 3/3) lorsqu'une étape échouée rend inaccessibles les étapes subséquentes. Les réponses ordonnées permettent de rendre compte de la progression des performances et donc des apprentissages. Parmi les modèles à réponses ordonnées, plusieurs approches sont considérées. Une première, l'approche cumulative, consiste à considérer chaque réponse comme un seuil possible pour un item dichotomique (Samejima, 1997). L'approche séquentielle consiste à considérer plutôt chaque réponse (ou étape) comme un item dichotomique autonome (Tutz, 1997). Enfin, une troisième possibilité, l'approche relative, consiste à utiliser deux réponses adjacentes pour former un item dichotomique (Masters & Wright, 1997). Cependant, comme cette dernière approche ne permet pas de fixer à priori l'ordre des réponses, elle convient moins au contexte édumétrique.

Avec les distributions habituelles (logistique ou normale), les approches énumérées précédemment sont incompatibles et/ou incohérentes. Cependant, selon Tutz (1991) de même que Bechger & Akkermans (2001), on peut démontrer que la distribution des valeurs extrêmes de Gumbel est la seule qui rend équivalentes les approches cumulative et séquentielle. Cette dernière distribution permet ainsi de profiter des propriétés fondamentales et intéressantes caractérisant chacune de ces approches. Parmi ces propriétés, on peut mentionner l'additivité et la divisibilité qui permettent de fixer librement le seuil de réussite d'une tâche polychotomique sans en modifier la modélisation, l'ordonnancement des points modaux toujours cohérent avec celui des réponses, et la prise en compte explicite des réponses absentes associées aux tâches séquentiellement inaccessibles. Enfin, la distribution des valeurs extrêmes est aussi la seule qui rend la mesure des apprentissages invariante par changement d'échelle quand on considère des tâches uniques ou des combinaisons de tâches pour évaluer les apprentissages.

Tel que présenté lors de communications précédentes (Riopel, 2017; Riopel *et al.*, 2017), des considérations édumétriques liées à l'invariance d'échelle permettent de proposer un modèle plus circonscrit et mieux adapté à la mesure des apprentissages en contexte scolaire. Évidemment, celui-ci permet d'abord de suivre les performances des participants, mais en y ajoutant la possibilité de déduire logiquement la forme mathématique des courbes observées pour l'apprentissage et la rétention. Mathématiquement, ce nouveau modèle repose explicitement sur trois hypothèses.

La première hypothèse, l'invariance de composition (ou de regroupement), propose que les lois liées à la performance ou à l'apprentissage sont les mêmes, à une constante d'échelle près, quand on combine des tâches équivalentes et indépendantes. Cette première hypothèse permet d'abord d'obtenir, à l'aide du théorème des valeurs extrêmes (Basrak, 2011), une fonction caractéristique qui prend la forme suivante :

$$P = \exp\{-\exp[-a\,(\theta - b)]\}\ ,$$

où P est la probabilité de réussite, θ est l'habileté du sujet, a est le paramètre de discrimination et b le paramètre de difficulté. La première hypothèse permet aussi de déduire une nouvelle loi explicite liant la probabilité de réussir au temps de réponse. Cette loi, présentée dans Riopel *et al.* (2017), peut être exprimée sous la forme suivante :

$$T = T_{min} - S \cdot \log(P)\ ,$$

où T est le temps pour réussir, T_{min} est le temps minimal (associé à la meilleure performance possible), S est une constante d'échelle et P est le taux de réussite associé à T.

La seconde hypothèse, l'invariance ordinale (ou de rang), propose que les lois sont les mêmes, à une constante d'échelle près, quand on attribue des rangs à des tâches ou à des groupes de tâches répétées. Cette seconde hypothèse permet de déduire la forme de la courbe régulière d'apprentissage (ou law of practice, dans la littérature anglo-saxonne, voir Heatcote, Brown et Mewhort, 2000). Cette loi, présentée dans Riopel (2017), peut être exprimée sous la forme suivante :

$$\log P = \log P_0 \cdot \exp(-\,a\,n^{c})\ ,$$

où P_0 est la probabilité initiale de réussir (quand n=0), a et c sont des constantes positives, et n correspond au nombre de répétitions de la tâche (ou du groupe de tâche).

La troisième et dernière hypothèse est l'invariance temporelle qui propose que les lois sont les mêmes, à une constante d'échelle près, quand on modifie uniformément la vitesse d'une séquence de tâches répétées. Cette hypothèse a déjà été proposée par Maylor *et al.* (2001). Elle permet de déduire la forme de la courbe de rétention ou d'oubli (ou forgetting curve, dans la littérature anglo-saxonne, voir Averell et Heatcote, 2011). Cette loi, présentée dans Riopel (2017), peut être exprimée sous la forme suivante :

$$\log P = \log P_0 \cdot \exp(- a\, t^{-b}),$$

où P_0 est la probabilité minimale de réussir (quand t tend vers l'infini), a et b sont des constantes positives, et t correspond à la durée de rétention.

C'est dans le contexte de ce nouveau modèle qu'une expérimentation a été effectuée pour évaluer des séquences d'identification de photographies d'oiseaux et de constellations à l'aide d'une application en ligne librement distribuée. Durant les deux dernières années, plus de deux mille participantes et participants, francophones et anglophones, ont accepté de participer à ce projet de recherche. La tâche consistait à observer une image puis à choisir la bonne identification parmi une liste proposée. Après chaque réponse donnée, le bon choix était surligné en vert durant trois secondes afin de favoriser un apprentissage. Les questions, en ordre aléatoire, étaient répétées plusieurs fois avec différentes échelles temporelles. Pour un participant donné, une expérimentation complète consistait en trois blocs indépendants de 129 questions, séparés par des pauses variables, pour un total de 387 questions et une durée totale d'environ 50 minutes en excluant les pauses. Les réponses données ainsi que les temps de réponses étaient mesurés. Les hypothèses d'invariance d'échelles ont d'abord été vérifiées selon une procédure générale en deux étapes qui ne fait pas d'hypothèse sur la forme attendue des courbes (Riopel *et al.*, 2017). Dans un premier temps, un paramètre d'échelle a été déterminé afin de rendre chaque paire de courbes liées aussi semblables que possible en minimisant la statistique du Chi-deux. Ensuite, cette même statistique a été utilisée dans un test du Chi-deux pour déterminer si la première courbe pouvait servir de modèle à la seconde. Les résultats obtenus jusqu'à présent permettent de confirmer les hypothèses d'invariance d'échelle et ainsi l'applicabilité du modèle. De plus, certaines propriétés du modèle proposé permettent de simplifier les calculs nécessaires à son application en contexte réel afin d'identifier des séquences optimales d'apprentissage.

Finalement, bien que le modèle proposé permette d'unifier plusieurs approches et conduise à de nouvelles lois prédictives et quantitatives pour les apprentissages, ses fondements théoriques explicites permettent de déduire quelques principes généraux d'analyse susceptibles d'enrichir aussi les interprétations qualitatives, soit 1) les tâches proposées aux apprenants peuvent toujours théoriquement être divisées ou combinées (le modèle s'appliquant indifféremment à toutes les échelles); 2) une erreur est toujours petite si on prend la peine de la mesurer précisément (même si ses conséquences en cascade peuvent être très grandes); 3) c'est en analysant les erreurs (plutôt que les succès) qu'on obtient le plus d'informations sur le niveau d'habileté des apprenants. Des travaux futurs pourraient permettre d'explorer les conséquences de ces lois et de ces principes dans une perspective édumétrique.

Mots-clés

Édumétrie, Invariance d'échelle, Courbe d'apprentissage, Courbe de rétention

Références bibliographiques

Averell, L., Heathcote, A. (2011). The form of forgetting curve and the fate of memories. *Journal of Mathematical Psychology, 55*(1), 25–35.

Basrak, B. (2011). Fisher-Tippett Theorem. Dans M. Lovric (dir.) *International Encyclopedia of Statistical Science* (p. 525-526). Berlin, Heidelberg: Springer Berlin Heidelberg.

Bechger, T.M. & Akkermans, W. (2001). A note on the equivalence of the graded response model and the sequential model. *Psychometrika 66*(3), 461-463.

De Landsheere, V. (1988). Faire réussir, faire échouer. *La compétence minimale et son évaluation.* Paris : Presses universitaires de France.

Heathcote, A., Brown, S., Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review, 7*(2), 185-207.

Masters G.N., Wright B.D. (1997). The Partial Credit Model. Dans van der Linden W.J., Hambleton R.K. (dir.) *Handbook of Modern Item Response Theory*. Springer, New York, NY

Maylor, E.A., Chater, N. & Brown, G.D.A. (2001). Scale invariance in the retrieval of retrospective and prospective memories. *Psychonomic Bulletin & Review, 8*(1), 162-167.

Riopel, M. (2017). Practice and forgetting curves deduced from scale invariance. *EDULEARN17 proceedings,* 4092-4096.

Riopel, M., Chastenay, P., Fortin-Clément, G., Potvin, P., Masson, S., Charland, P. (2017). Using invariance to model practice, forgetting and spacing effect. *EDULEARN17 proceedings,* 4334-4341.

Samejima F. (1997). Graded Response Model. Dans van der Linden W.J., Hambleton R.K. (dir.) *Handbook of Modern Item Response Theory.* Springer, New York, NYTutz (1997)

Tutz, G. (1991). Sequential models in categorical regression. *Computational Statistics & Data Analysis, 11*, 275–295.