How Success in a Task Depends on the Skills Level: Two Uncertainty-Based Justifications of a Semi-Heuristic Rasch Model

Joe Lorkowski[1], Olga Kosheleva[2], and Vladik Kreinovich[1]

Departments of [1]Computer Science and [2]Teacher Education
University of Texas at El Paso
500 W. University
El Paso, Texas 79968, USA
lorkowski@computer.org, olgak@utep.edu, vladik@utep.edu

1. An Empirically Successful Rasch Model

- For each level of student skills, the student is usually:
  - very successful in solving simple problems,
  - not yet successful in solving problems which are – to this student – too complex, and
  - reasonably successful in solving problems which are of the right complexity.
- To design adequate tests, it is desirable to understand how a success s in a task depends:
  - on the student's skill level ℓ and
  - on the problem's complexity c.
- Empirical *Rasch model* predicts $s = \dfrac{1}{1 + \exp(c - \ell)}$.
- Practitioners, however, are somewhat reluctant to use this formula, since it lacks a deeper justification.

2. What We Do

- In this talk, we provide two possible justifications for the Rasch model.
- The first is a simple fuzzy-based justification which provides a good intuitive explanation for this model.
- This will hopefully enhance its use in teaching practice.
- The second is a somewhat more sophisticated explanation which is:
  - less intuitive but
  - provides a quantitative justification.

3. First Justification for the Rasch Model

- Let us fix c and consider the dependence $s = g(\ell)$.
- When we change ℓ slightly, to $\ell + \Delta\ell$, the success also changes slightly: $g(\ell + \Delta\ell) \approx g(\ell)$.
- Thus, once we know $g(\ell)$, it is convenient to store not $g(\ell+\Delta\ell)$, but the difference $g(\ell+\Delta\ell) - g(\ell) \approx \dfrac{dg}{d\ell} \cdot \Delta\ell$.
- Here, $\dfrac{dg}{d\ell}$ depends on $s = g(\ell)$: $\dfrac{dg}{d\ell} = f(s) = f(g(\ell))$.
- In the absence of skills, when $\ell \approx -\infty$ and $s \approx 0$, adding a little skills does not help much, so $f(s) \approx 0$.
- For almost perfect skills $\ell \approx +\infty$ and $s \approx 1$, similarly $f(s) \approx 0$.
- So, $f(s)$ is big when s is big ($s \gg 0$) but not too big ($1 - s \gg 0$).

4. First Justification for the Rasch Model (cont-d)

- Rule: $f(s)$ is big when:
  - s is big ($s \gg 0$) but
  - not too big ($1 - s \gg 0$).
- Here, "but" means "and", the simplest "and" is the product.
- The simplest membership function for "big" is $\mu_{\text{big}}(s) = s$.
- Thus, the degree to which $f(s)$ is big is equal to

$$s \cdot (1 - s) : \quad f(s) = s \cdot (1 - s).$$

- The equation $\dfrac{dg}{d\ell} = g \cdot (1 - g)$ leads exactly to Rasch's model $g(\ell) = \dfrac{1}{1 + \exp(c - \ell)}$ for some c.

5. What If Use min for "and"?

- What if we use a different "and"-operation, for example, $\min(a, b)$?
- Let us show that in this case, we also get a meaningful model.
- Indeed, in this case, the corresponding equation takes the form $\dfrac{dg}{d\ell} = \min(g, 1 - g)$.
- Its solution is:
  - $g(\ell) = C_- \cdot \exp(\ell)$ when $s = g(\ell) \le 0.5$, and
  - $g(\ell) = 1 - C_+ \cdot \exp(-\ell)$ when $s = g(\ell) \ge 0.5$.
- In particular, for $C_- = 0.5$, we get a cdf of the Laplace distribution $\rho(x) = \dfrac{1}{2} \cdot \exp(-|x|)$.
- This distribution is used in many applications – e.g., to modify the data in large databases to promote privacy.

6. Towards a Second Justification

- The success s depends on how much the skills level ℓ exceeds the complexity c of the task: $s = h(\ell - c)$.
- For each c, we can use the value $h(\ell - c)$ to gauge the students' skills.
- For different c, we get different scales for measuring skills.
- This is similar to having different scales in physics:
  - a change in a measuring unit leads to $x' = a \cdot x$; e.g., 2 m $= 100 \cdot 2$ cm;
  - a change in a starting point leads to $x' = x + b$; e.g., 20° C $= (20 + 273)^\circ$ K.
- In physics, re-scaling is usually linear, but here, $0 \to 0$, $1 \to 1$, so we need a non-linear re-scaling.

7. How to Describe Not-Necessarily-Linear Re-Scalings

- If we first apply one reasonable re-scaling, and after that another one, we still get a reasonable re-scaling.
- For example, we can first change meters to centimeters, and then replace centimeters with inches.
- Then, the resulting re-scaling from meters to inches is still a linear transformation.
- In mathematical terms, this means that the class of reasonable e-scalings is closed under composition.
- Also, if we have a re-scaling, e.g., from C to F, then the "inverse" re-scaling from F to C is also reasonable.
- In precise terms, this means that the class of all reasonable re-scalings is invariant under taking the inversion.

8. How to Describe Re-Scalings (cont-d)

- Thus, we can say that reasonable re-scalings form a transformation group.
- Our goal is computations.
- In a computer, we can only store finitely many parameters.
- Thus, each re-scaling must be determined by finitely many parameters.
- Such groups are called *finite-dimensional.*
- So, we need to describe all finite-dimensional transformation groups that contain all linear transformations.
- It is known that all functions from these groups are fractionally-linear $f(s) = \dfrac{a \cdot s + b}{c \cdot s + d}$.

9. Resulting Equation

- We consider a transformation $s' = f(s)$ between
$$s = h(\ell - c) \text{ and } s' = h(\ell - c').$$
- We showed that this transformation is fractionally-linear $f(s) = \dfrac{a \cdot s + b}{c \cdot s + d}$.
- When $s = 0$, we should have $s' = 0$, hence $b = 0$.
- We can now divide both numerator and denominator by d, then $f(s) = \dfrac{A \cdot s}{C \cdot s + 1}$.
- When $s = 1$, we should have $s' = 1$, so $A = C + 1$, and $f(s) = \dfrac{(1 + C) \cdot s}{C \cdot s + 1}$.
- For $c' = 0$, we thus get
$$h(\ell - c) = \frac{(1 + C(c)) \cdot h(\ell)}{C(c) \cdot h(\ell) + 1}.$$

10. Solving the Resulting Equation Explains the Rasch Model

- We know that

$$h(\ell - c) = \frac{(1 + C(c)) \cdot h(\ell)}{C(c) \cdot h(\ell) + 1}.$$

- Differentiating both sides w.r.t. c and taking $c = 0$, we get a differential equation whose general solution is

$$h(\ell) = \frac{1}{1 + \exp(k \cdot (c - \ell))}.$$

- By changing measuring units for ℓ and c to k times smaller ones, we get the Rasch model

$$h(\ell) = \frac{1}{1 + \exp(c - \ell)}.$$

11. Conclusion

- It has been empirically shown that,
  - once we know the complexity c of a task, and the skill level ℓ of a student attempting this task,
  - the student's success s is determined by Rasch's formula

$$s = \frac{1}{1 + \exp(c - \ell)}.$$

- In this talk, we provide two uncertainty-based justifications for this model:
  - a simpler fuzzy-based justification provides an intuitive semi-qualitative explanation for this formula;
  - a more complex justification provides a quantitative explanation for the Rasch model.

12. Acknowledgments

- This work was supported in part by the National Science Foundation grants:
  - HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and
  - DUE-0926721.
- The authors are greatly thankful to the anonymous referees for valuable suggestions.