NEWTON:
PINBALL WIZARD?

Larry H. Ludlow[1]
Kathleen A. Haley
Boston College

In 1981 the first author worked as Geoff Master's MESA Psychometric Laboratory assistant. His task was to write code for the partial credit Rasch model. While testing the program a computational problem was encountered that did not make sense, regardless of how, or how often, Geoff or Ben tried to explain it. The problem was that some of the calculations were shrinking to zero and the program was crashing. Geoff and Ben explained this phenomenon as a failure to converge on the part of the Newton-Raphson technique. That explanation did not mean much to him then nor does it mean much now to many of our beginning psychometrics students.

In these days of specialization, it is amazing and humbling to reflect on the great number of areas in which Newton not only had interest, but great influence. Optics, astronomy, chemistry, physics, and mathematics are notable examples. An Alexander Pope couplet is often quoted to demonstrate Newton's influence in his lifetime:

Nature, and Nature's Laws lay hid in Night.
God said, Let Newton be! and All was Light.

Newton's beginnings, however, were not auspicious. He was born dangerously prematurely and was lucky to survive. His father died before he was born, and his mother was absent for much of his early childhood, having left him with his grandmother when she remarried (Christianson, 1984).

Although his work changed science in ways we rarely think about, one of his greatest contributions to our field was in pure mathematics. His work in the invention of differential calculus allows us to find minima and maxima of curves, without which we would be missing many standard statistical techniques. In addition, while trying to solve Kepler's equation for the position of a planet at any given time he developed a numerical procedure for solving higher-order polynomials that could not be solved using calculus (Pepper, 1988).

The rationale behind the technique is relatively simple. That is, if we don't know how to obtain a direct estimate of a parameter, then use a rough initial estimate that we can iteratively adjust. Different numerical analysis strategies exist for obtaining initial estimates, their subsequent adjustments, and the final estimates. Newton's contribution was that he developed the first practical technique for solving such a problem. The technical detail is standard material in numerical analysis texts and is included in the appendix.

The way the original programming problem came to be understood, however, was by printing out all the intermediate values of the calculations and then plotting them.

Figure 1

Newton-Raphson Divergence

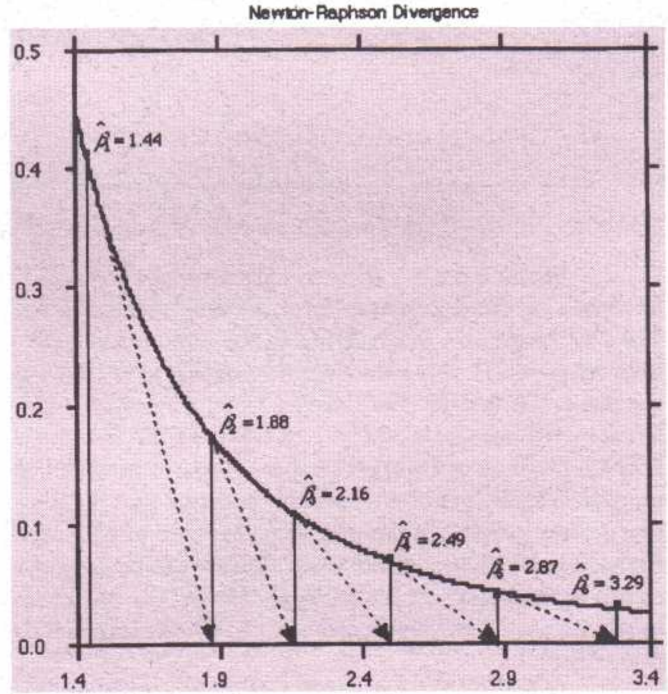

Figure 1 illustrates the situation that prompted the study of the Newton-Raphson technique. In this situation we see divergence, or a failure to converge. The initial, intermediate, and final calculations **(see Table 1)** in the iterative process to arrive at the logit ability estimate for an extremely high raw score are plotted. The estimates ($\hat{\beta}$) become more extreme as the part of the equation that reflects the precision of the estimates

(DB) goes to zero causing the actual adjustment (DINC) to increase without limit. This figure did not offer particularly useful insight into the technique, however, because it does not show what a proper solution should look like. (NOTE: the terms GS, DB, and DINC were variable names originally used in the programs SCALE and CREDIT and may very well still be used in whatever Rasch software you are currently using. In addition, most programs have built-in checks that slow down DB from going to zero—the "1/2 Correction" factor applied to the adjustment.)

Figure 2
Newton-Raphson Convergence

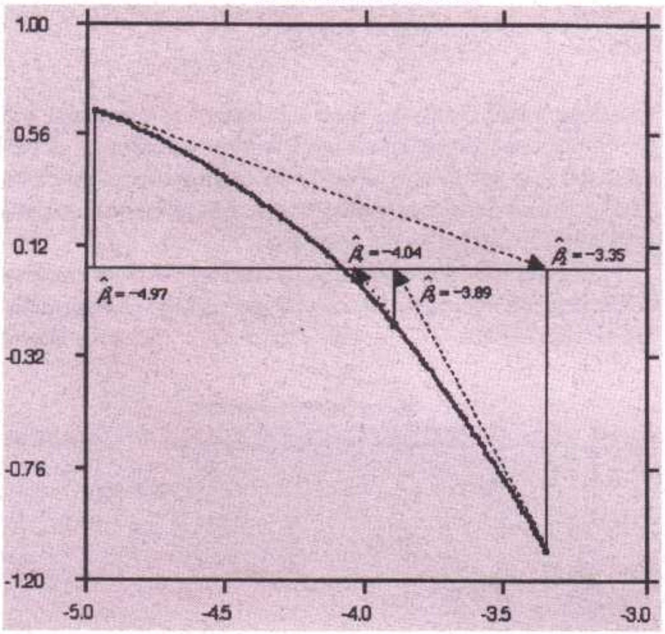

Figure 2 is a plot of the calculations **(see Table 2)** to arrive at a logit ability estimate for an extremely low raw score. The plot illustrates a convergence in the estimates. The first estimate is (-4.97). The second estimate swings to the right on the line and is (-3.35). The third estimate swings back to the left along the line and is (-3.89). Note also how the adjustment (DINC) becomes increasingly smaller. After studying plot after plot of similar patterns it suddenly occurred that the iterative process (when it converges) is analogous to a ball bouncing back and forth in a pinball game until the ball finally comes to rest at the bottom of the machine. Hence, the title of the article (although converging estimates do not necessarily have to oscillate).

Ironically, Newton would not likely have found humor in the analogy or the title of this article. In fact, he was described as "anything but affable." He was extremely unwilling to give credit to others or share credit for great discoveries. For example, he conducted a near-lifelong priority dispute with Leibniz over the invention of calculus, and he had a long-standing dispute with Johann Bernoulli. He was acrimonious in dispute, vindictive, imperious, insulting, and arrogant. His "academic overkill" even went so far as to lead him to brag "He had broke Leibniz's Heart with his Reply to him" (Boorstein, 1983).

Finally, how did "Raphson" become associated with the technique? John Raphson was an accomplished mathematician in his own right, having been elected to the Royal Society the year *before* his graduation from Jesus College Cambridge. He published a mathematical dictionary and *De spatio reali*, an application of mathematical reasoning to theological issues. He also published another theological book, *Demonstratio de deo* (O'Connor & Robertson, 1998). However, he is better known by his association with Newton. Quite simply, the numerical technique developed by Newton in 1671 came to be called the Newton-Raphson technique because it was first published in Joseph Raphson's 1690 *Analysis aequationum universalis*. Newton himself did not publish it until 1736 (O'Connor & Robertson, 1998).

References

Boorstein, D.J. (1983). The Discoverers. New York: Random House.

Christianson, G. (1984). In the Presence of the Creator: Isaac Newton and His Times. New York: The Free Press, Macmillan.

Pepper, J. (1988). Newton's mathematical work. In Fauvel, J., Flood, R., Shortland, M., and Wilson, R. (eds.), Let Newton Be!. Oxford: Oxford University Press.

Appendix

The pinball analogy often works better in our courses than the traditional explanation found in numerical analysis texts: Given a function f(x), solve for the roots of f(x) such that f(x)=0. To attempt to find that location where the roots of the function are 0 requires some initial estimate for the location f(x)=0. Once an initial estimate is provided we try to improve upon it. The solution is based on the observation that if x_0 is close to a 0 of f(x), then the tangent to the graph of f(x) at $(x_0, f(x_0))$ intersects the axis X at a point, say $x^{(1)}$, which is closer than x_0 to the 0 of f. The process then consists of computing $x^{(1)}$, substituting it back into the equation as x_0, and re-computing $x^{(1)}$ until some level of desired accuracy is achieved. This process is the Newton-Raphson method of solving for the roots of an equation.

For our purposes the important piece in this process becomes $x^{(1)} = x_0 - \frac{f(x_0)}{f'(x_0)}$. The numerator (called GS) gives the direction in which the adjustment should be made and the denominator (called DB) reflects the magnitude of adjustment (DINC). Technically, the second derivative of the function must not go to zero (shown by the asymptotic function) which is exactly what happens with some data sets containing extremely high or low person or item scores.

In practical terms, when we are estimating β for persons the function f(x) we are solving (GS) is the difference $(r - \sum p)$ where r is a person total score and $\sum p$ corresponds to the estimated total score. The value $x^{(1)}$ corresponds to our new estimate of $\hat{\beta}$ and x_0 corresponds to our previous estimate of $\hat{\beta}$ (which on the first iteration came from PROX).

Table 1. Divergence

Iteration	$\hat{\beta}$	GS	DB	DINC	1/2Correction
1	1.44	.41	-.94	.43	
2	1.88	.17	-.3	.57	.28
3	2.16	.11	-.16	.67	.33
4	2.49	.07	-.09	.76	.38
5	2.87	.04	-.05	.83	.42
6	3.29	.03	-.03	1.86	.43

The following example shows how the equation works:

Given that $x^{(1)} = x_0 - \frac{f(x_0)}{f'(x_0)}$, then $\hat{\beta}_2 = \hat{\beta}_1 - \frac{GS}{DB}$, or

$1.88 = 1.44 - \frac{.41}{-.94}$ where $DINC = -\frac{GS}{DB}$ or .43.

Table 2. Convergence

Iteration	$\hat{\beta}$	GS	DB	DINC
1	-4.97	.65	-.4	1.62
2	-3.35	-1.09	-2.03	-.53
3	-3.89	-.2	-1.31	-.15
4	-4.04	-.01	-1.12	-.01

[1] This paper is distilled from one originally written in 1981 with the statistical assistance of Graham Douglas and Geofferey Masters. The original paper is still used in our courses.

Larry H. Ludlow, Ph.D.

Associate Professor, Boston College, School of Education, Education Research, Measurement, and Evaluation Program.

Professional interests: developing interesting graphical representations of multivariate data (visualizing an eigenvector), and applying psychometric models in situations where the results have an obvious practical utility (scaling flute performance).

Personal interests: woodcarving, sketching, and motorcycling.

Last book read: Arthur Koestler, The Sleepwalkers.

Personal goal: Actually catch something fly-fishing.

Favorite drink: Diet Dr. Pepper.

Favorite quote: "If it exists, it can be measured. If it can't be measured, it doesn't exist." (mine)

e-mail: LUDLOW@BC.EDU