Giving Notice: Using Cauchy Sequencing to Inform Diplomates of Longitudinal Assessment Progress

ABR | American Board of Radiology

As the first cohort of diplomates approaches their five-year performance evaluation for Online Longitudinal Assessment (OLA), it is important to notify those who might fail, so they can work to improve their score or plan to take an alternate assessment to maintain certification. Ideally, candidates should receive a notification in time to impact their future OLA performance prior to the five-year assessment deadline. Scoring is continuous and thus sequential, much like a Cauchy sequence. We use OLA assessment data from a variety of disciplines within radiology to demonstrate how certifying organizations can employ Cauchy sequencing to predict diplomates' pass/fail score decisions in a timely manner.

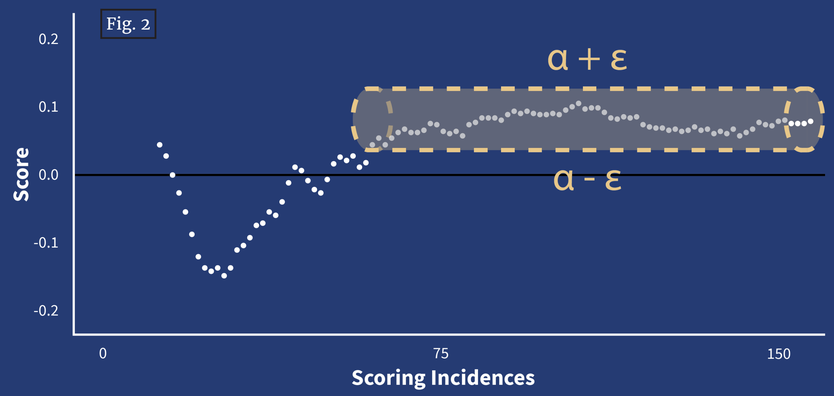
How It Works

Imagine a ball that has been dropped to the floor (Fig. 1). The first few times the ball bounces, it will bounce relatively highly. The longer it bounces, the shorter the bounces become. Eventually, the height of the bounces becomes negligible. In a Cauchy sequence, measurements at each interval become arbitrarily close together after sufficient measurement observations.

Fig. 1

Now imagine that ball is one person's daily score (Fig. 2). At the beginning of OLA, their score bounces up and down; however, over time, it begins to flatten.

Fig. 2

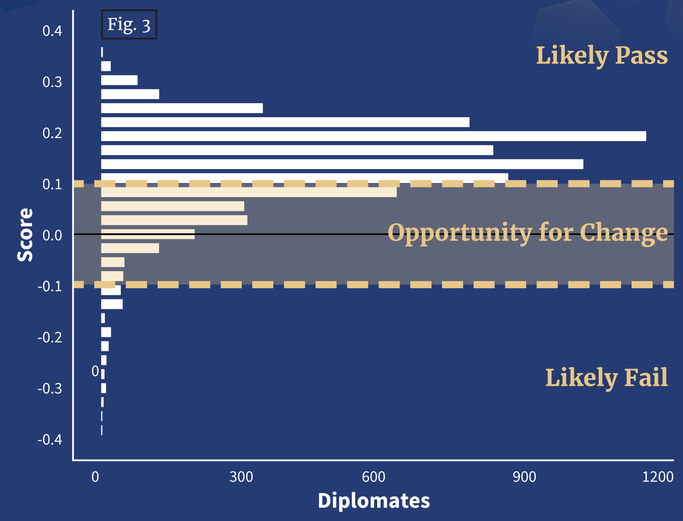


This flattening visually depicts the convergence into a Cauchy sequence. Each dot is one person's score that appears at different intervals of time. In the diagram, epsilon (**ε**) represents a measure of closeness, while alpha (**α**) is the average score over time. Epsilon equals maximum score minus the mean score within the sequence. The epsilon tube (shown in orange), shows the overall range of closeness between the scores. Convergence occurs when this range becomes sufficiently small such that a Cauchy sequence is detected.

How It Helps Us

The inherent properties of Cauchy sequences allow us to make predictions at various levels of accuracy. In application, the number of longitudinal scores needed for convergence is critical. If the minimum number of scores needed for convergence is too high, diplomates cannot be notified early enough to take action to improve their performance. If the minimum is low, predictions can be made earlier, giving diplomates more information about their OLA progress and how it will impact Part 3 of Continuing Certification.

Fig. 3



Research Questions

RQ1: Does convergence happen with the number of scores available at the beginning of the fourth year of participation (156 scoring incidences)?
RQ2: Is convergence robust to sample size?

(RQ1) The fifth year is when performance is assessed for the purposes of Continuing Certification. Therefore, we need to know if the minimum number of items completed at the beginning of the fourth year is sufficient for convergence. Convergence happened at 156 scoring incidences.

Sample Size (n)	Repetitions (k)	Epsilon Tube	Likely End Score Pass/Fail Change
10,000	*	+/- 0.08	1.40%
5,000	10	+/- 0.08	1.50%
1,000	10	+/- 0.12	1.90%
500	10	+/- 0.13	2.50%
100	10	+/- 0.18	1.70%

Notes: Sampling with replacement.
No repetitions were completed for n = 10,000 due to insufficient data.
Tests of sample sizes 5,000 and below were repeated 10 times to check for anomalous findings. All observations had at least 156 scoring incidences.

(RQ2) To test for robustness to sample size, the same group was sampled with replacement at sample sizes ranging from 10,000 to 100 individuals. Each size grouping was randomly sampled 10 times. As shown in the table, Cauchy sequencing occurred at all samples sizes; larger sample sizes yielded smaller epsilon tubes.

How It Could Help You

For another ABMS board to use these methods, several considerations around scoring should be reviewed. First, the American Board of Radiology places OLA scores on a z-scale, where zero and above are passing. It is important to determine if a different scoring scale might yield Cauchy sequences.

More important, the sequences within our data are strongly related to the number of items a diplomate answers over an extended time. Scoring is calculated weekly by the ABR, yielding 52 observations per year per diplomate, because new items are released every week. Through the minimum participation requirements, every diplomate generates at least 156 scoring incidents by the end of the third year (though many have more by then).

For sequencing to occur, scoring must be calculated over time. If diplomates answered 104 items during a single day, which is not permissible in OLA, the longitudinal aspects of the data would be lost. There would only be one scoring incident for that year, rather than 104. OLA must be continuous to maintain longitudinal data properties.

Other ABMS boards should consider their own scoring practices when determining if this methodology might yield useful predictions.