J. Math. Sci. Model., 9(2) (2026) 107-119 https://doi.org/10.33187/jmsm.1809283

Journal of Mathematical Sciences and Modelling

ISSN (Online) 2636-8692
https://dergipark.org.tr/en/pub/jmsm

Research Article

AI-Supported Interactive Simulation to Teach Statistical Hypothesis Testing: Case of the Point Optimal Test and Power Envelope for the Cauchy Distribution

Asad Ul Islam Khan[1], Mehmet Akın Bulut[2] and Adem Yurdunkulu[2]*

[1]*Economics Department, Ibn Haldun University, Istanbul, Türkiye*
[2]*School of Education, Education Sciences, Ibn Haldun University, Istanbul, Türkiye*

Article Info

Keywords: *AI-supported learning, Interactive simulation-based learning, Power envelope, Point optimal test, Power, Size*

Received: *23 October 2025*
Accepted: *11 May 2026*
Available online: *12 May 2026*

Abstract

Teaching complex statistical concepts like the Neyman–Pearson Lemma and Point Optimal testing is often hindered by their mathematical abstraction. To facilitate the learning process, education technology researchers offer interactive, simulation-based, technology-enhanced teaching solutions. Thus, this paper presents an AI-supported interactive simulation module, integrated into a university Learning Management System (Canvas LMS), to bridge the gap between rigour and comprehension. Using the location parameter of a Cauchy distribution as a case study, it is demonstrated how students can dynamically visualize shifting rejection regions and power envelopes in real time. The module employs a reinforcement-learning-based engine to analyse learner interactions, providing adaptive hints that address specific misconceptions regarding distributional symmetry and test efficiency. Preliminary implementation in advanced econometrics courses indicates that this interactive, AI-driven learning and teaching approach significantly improves conceptual understanding and student engagement by transforming static theory into an exploratory learning experience.

1. Introduction and Literature Review

The two papers by Durbin and Watson [1,2] are among the highly cited econometrics articles of that era. These papers were on "testing of serial correlation". Back then, it was a well-established fact that time series has a serious problem of serial correlation, and it is widespread, but all of the tests for serial correlation were dependent on the regressors' matrix structure. Durbin and Watson tabulated the critical values which were invariant to the regressors' matrix structure. At that time, without this table, it was impossible to test for serial correlation due to the limitations in computational capabilities. Recently, the advances in computational capabilities have transformed the task of obtaining simulated critical values for any statistic into a trivial matter. In the current era, journals may even reject the Durbin–Watson paper as it provides an unnecessary approximation.

In the same way, the testing of hypothesis procedure was based on asymptotic theory and a table of critical values was developed on the basis of asymptotic approximations. These critical values (asymptotic ones) were used even for finite and small samples and the use of these asymptotic critical values might lead to wrong conclusions. This implies that symmetry is considered in the use of asymptotic critical values. Whether the sample size is small and finite or large or infinite, the same asymptotic critical values will be considered. However, in practice these asymptotic critical values are asymmetric to finite and small sample sizes as compared to large and/or infinite sample sizes. In the literature, the use of asymptotic critical values is a common practice. There are many studies citing the fact that the use of asymptotic critical values leads to wrong conclusions and erroneous policy recommendations [3–5]. Due to computational limitations, it was impossible to get the critical values for finite and small samples [6]. However, owing to computational advancement, it is now possible to compute the test statistic and find its simulated critical values for finite samples and asymptotic approximations are no longer required.

* *Corresponding author*
***Email addresses*: asad.khan@ihu.edu.tr, akin.bulut@ihu.edu.tr, adem.yurdunkulu@ihu.edu.tr**
Cite as: A. U. I. Khan, M. A. Bulut, and A. Yurdunkulu, *AI-supported interactive simulation to teach statistical hypothesis testing: Case of the point optimal test and power envelope for the Cauchy distribution*, J. Math. Sci. Model., 9(2) (2026), 107–119.

Neyman and Pearson [7] provided the Neyman–Pearson Lemma and defined the point optimal test at a specific alternative. Before the computational advancement, this lemma was applied to a very limited class of problems due to computational inadequacy. Now, due to computational advancement, researchers, statisticians and practitioners can use this lemma to find the point optimal test at every specific alternative. Moreover, the simulated critical values can be computed (although this will be computationally burdensome).

As far as the comparison of tests is concerned, tests were compared using asymptotic theory, but this was possible for only a limited class of problems [8]. Outside this class, the problem of finding a better test than others was not only impossible but "even without formulation" [8]. Lehmann and Romano [8] also explained the "Stringency criterion", a robust criterion for test comparison considering the whole alternative space. For the stringency criterion, the point optimal test is needed at every specific alternative. For detailed discussion on stringency and its use in test comparison, see [4–6, 9, 10].

The theory of hypothesis testing is a fundamental component of econometrics, yet its presentation in most texts remains highly abstract, often leaving students with a procedural rather than conceptual understanding [11]. Concepts such as the Neyman–Pearson Lemma, point optimal tests, and power envelopes are mathematically elegant but require a high level of statistical maturity to fully grasp. Traditional lecture-based approaches may fail to convey the intuition behind these ideas, particularly in distinguishing the behaviour of tests under symmetric and asymmetric distributions.

In parallel, advances in educational technology and artificial intelligence in education (AIED) have opened new possibilities for teaching quantitative subjects. Simulation-based tools allow learners to manipulate parameters, visualize distributions, and observe real-time changes in statistical properties [12]. When enhanced with AI-driven adaptivity, such systems can monitor student interactions, detect patterns of misunderstanding, and provide personalized feedback [13]. This aligns with the shift toward active, data-driven learning environments in higher education, particularly in statistical disciplines.

This paper combines rigorous econometric theory with AI-supported interactive simulation to enhance the teaching and learning of hypothesis testing. Using the location parameter of a single draw from a Cauchy distribution as a case study, the development of the point optimal test and its associated power envelope are illustrated. The accompanying AI-enhanced module, integrated into the university's Learning Management System, enables students to explore these concepts dynamically, thereby bridging the gap between abstract mathematical rigour and learner comprehension. The AI-supported module is designed around several core pedagogical objectives. By engaging with the interactive simulation, students are expected to achieve the following learning outcomes:

- **To conceptualize power and size trade-offs:** Develop an intuitive understanding of the inverse relationship between Type I error (α) and Type II error (β), and how the power of a test $(1-\beta)$ is affected by changing significance levels.
- **To interpret likelihood ratios dynamically:** Move beyond formulaic calculation to understand the likelihood ratio as a measure of how strongly an observed outcome favours the alternative hypothesis over the null.
- **To identify distributional symmetry and asymmetry:** Visualize how the shape of the rejection region (symmetric vs. asymmetric) and the resulting p-values shift depending on the underlying distribution and the specific alternative hypothesis being tested.
- **To master point optimal and power envelope theory:** Grasp the "point optimal" nature of the Neyman–Pearson test by observing its performance at specific points in the alternative space and understanding the power envelope as the theoretical upper bound for test comparison.

Table 1.1 displays the symbols, definitions, and interpretations used throughout the study.

Symbol	Definition / Interpretation
X, x	Random variable and its observed realisation
θ	Location parameter of the Cauchy distribution
θ_0	Value of θ under the null hypothesis; $\theta_0 = 0$
θ_1	Value of θ under the point alternative; $\theta_1 > 0$
θ^*	Critical threshold; $\theta^* = 2/k$
H_0	Null hypothesis: $\theta = \theta_0 = 0$
H_1	Alternative hypothesis: $\theta = \theta_1 > 0$
$l(x,\theta)$	Likelihood function (density) of x at parameter θ
$LR(x,\theta_1,\theta_0)$	Likelihood ratio: $l(x,\theta_0)/l(x,\theta_1)$
c	Critical constant in the NP rejection rule; $LR \leq c$ defines $R^*(c)$
$R^*(c)$	Optimal (point optimal) rejection region for constant c
r_1, r_2	Roots of the quadratic inequality defining boundaries of $R^*(c)$
α	Nominal significance level (probability of Type I error)
β	Probability of Type II error
$\pi_\alpha(\theta_1)$	Power envelope: maximum power at θ_1 for any test of level α
k	Shorthand for $\tan(\pi\alpha)$

Table 1.1: List of symbols

As for the literature analysis on education technology and artificial intelligence in education, statistical hypothesis testing, particularly the Neyman–Pearson framework, remains a cornerstone of econometrics, yet its abstract nature poses significant pedagogical challenges [11, 14]. Concepts such as point optimal tests and power envelopes, though theoretically elegant, are difficult for students to internalize without dynamic, visual exploration [15]. Educational technology has been shown to bridge this gap, with simulation-based learning environments enabling learners to manipulate statistical parameters and observe immediate changes in probability distributions and test performance [12]. When integrated into Learning Management Systems, these tools can provide a seamless, course-embedded experience that supports both

computation and visualization [16]. Recent advances in AI in education extend this capability through adaptive learning systems, which monitor learner interactions, detect misconceptions, and deliver personalized feedback in real time [13, 17]. In mathematics and statistics education, such AI-driven approaches have been found to improve both conceptual understanding and retention by tailoring the instructional pathway to the learner's evolving needs [18]. By combining rigorous econometric theory with AI-supported, interactive simulation, it is possible to create a learning environment where abstract concepts like the power envelope become tangible, testable, and personally relevant to each student.

The landscape of higher education has seen a shift where artificial intelligence (AI) has transitioned from a niche computing domain to an infrastructural presence across diverse disciplines, including business, law, teacher education, and the humanities. However, AI is not a stable or universally agreed construct; existing frameworks variously combine technical understanding, operational competence, and ethical awareness, often without clearly articulating how these dimensions relate to one another [19, 20].

Much of the emerging literature treats AI as a competency-based outcome, yet these studies frequently rely on self-reported measures and short-term evaluations. As a result, AI is often operationalized as perceived readiness rather than demonstrated transformation in pedagogical reasoning or epistemic practice. Research suggests that conversational agents can enhance motivation and perceived instructional support, particularly when embedded within structured activities rather than deployed as autonomous substitutes for teachers [21, 22]. More recent analyses of generative tools like LMS point to shifts in assessment design and feedback practices while exposing tensions related to authorship and academic integrity [23].

Scholarship emphasizes that learning outcomes are strengthened when students actively manipulate representations, test ideas, and engage in dialogic reflection rather than passively consuming content [24, 25]. In the context of the statistical simulation:

- AI contributes most productively when embedded within iterative, inquiry-oriented activities that require interpretation, critique, and revision [26, 27].
- Adoption of tools like ChatGPT is shaped by perceived usefulness and ease of use, alongside concerns about reliability and epistemic accountability [28].
- Recent contributions call for design-oriented frameworks that support educators in orchestrating AI tools within disciplinary, institutional, and ethical constraints [29, 30].

By analysing the AI-supported simulation through this lens, the gap between technological inclusion and pedagogical configuration is addressed. The analysis examines how specific design decisions, such as the reinforcement-learning-based adaptive feedback, shape participants' practices and professional orientations within AI-mediated learning environments.

2. Methodology

To extend the theoretical framework into an educational technology context, an AI-enhanced, web-based simulation was developed and integrated into the university's Learning Management System (LMS). The simulation was built using Python for the statistical computations (including the point optimal test and power envelope for the Cauchy distribution) and JavaScript/D3.js for interactive visualizations. Students can adjust key parameters (e.g., significance level, location parameter shift, sample size) and immediately observe changes to the test statistic distribution and the associated power curve.

A reinforcement-learning-based recommendation engine within the module monitors student interactions, identifying patterns of repeated errors or misinterpretations. Based on these patterns, the system delivers targeted feedback, visual explanations, and scaffolded practice problems. The tool logs engagement data (time on task, sequence of parameter changes, number of simulation runs) to provide instructors with dashboards for real-time monitoring of student progress. This design allows abstract hypothesis-testing theory to be experienced as an interactive, exploratory process rather than a static, lecture-only concept, thereby increasing learner engagement and comprehension.

2.1. Mapping theoretical components to interactive features

To ensure that the mathematical rigour of the Cauchy location parameter test is translated into conceptual understanding, each theoretical component is mapped to a specific functional element within the LMS module:

- **Rejection region shapes (R_c^*):** The simulation dynamically renders the three possible shapes of the rejection region — the single interval ($c < 1$), the one-sided ray ($c = 1$), and the union of two intervals ($c > 1$) — as students manipulate the alternative hypothesis parameter θ_1.
- **The critical threshold (θ^*):** The theoretical value $\theta^* \approx 12.6275$ (for $\alpha = 0.05$) acts as a "phase transition" point. As students slide θ_1 across this threshold, the visualization provides a real-time animation of the rejection region shifting from a bounded interval to an unbounded dual-interval.
- **Power envelope curvature ($\pi_\alpha(\theta_1)$):** While students can calculate the power for any specific test, the module overlays the theoretical power envelope — the maximum attainable power defined in Section 6. This allows students to visually assess the distance between their current test's performance and the best possible theoretical limit.
- **Likelihood ratio interpretation:** The module includes a "Point-Probe" feature where selecting a specific outcome x reveals the exact likelihood ratio $L(x, \theta_1, 0)$. This helps students internalize why points far in the tails may still favour the null hypothesis if they are "even more unlikely" under the alternative.

2.2. The AI adaptivity framework: RL-based feedback

The simulation's "personalized tutor" capability is driven by a reinforcement-learning policy that treats the student's mastery as the "state" and pedagogical interventions as "actions". The mechanism is structured as follows.

Error detection categories. The system logs interaction telemetry to identify three primary misconceptions: (i) *Symmetry bias* — a student consistently expects a two-tailed symmetric rejection region regardless of the θ_1 value; (ii) *Power-size confusion* — a student attempts to maximize power by increasing α without recognizing the resulting increase in Type I error; (iii) *Threshold ignorance* — failing to adjust the test structure when crossing the critical θ^* boundary.

Triggering logic. Feedback is triggered through a learned policy that evaluates "Engagement State" vs. "Accuracy". If a student performs five consecutive parameter changes that move the test further away from the power envelope without a change in α, the system identifies a "conceptual bottleneck".

Adaptive hinting examples.

- *Initial hint*: "Observe how the rejection region shifts as you increase θ_1. Is it still symmetric?"
- *Scaffolded prompt*: "You've reached $\theta^* \approx 12.63$. Look at the likelihood ratio at the tails — does a single interval still capture the most likely outcomes under H_1?"

As shown in the LMS Dashboard (Appendix A), the system uses real-time polling to calibrate its feedback. If a student rates their "Power Envelope Understanding" as low (1 or 2) in the embedded Poll, the AI increases the frequency of visual explanations over text-based hints. This data-driven approach allows the instructor to view a "Heatmap of Misconceptions" across the entire cohort, identifying if the majority of the class is struggling with the transition at θ^* or the trade-off depicted in Table 4.2.

Following the ed-tech literature review, a detailed review of the definitions and concepts about significance testing is provided in Section 3. Section 4 addresses the likelihood ratio and other tests. Section 5 lays out the specialised case of the Cauchy location parameter.

2.3. Technical architecture of the RL adaptive engine

The adaptive feedback mechanism is structured as a Knowledge-Tracing Reinforcement Learning (KTRL) model. The system treats the student's interaction sequence as a series of states in a Markov Decision Process (MDP) to optimize for the reward of conceptual mastery.

- **State space (S):** Defined by a vector of student inputs, including the current α level, the chosen θ_1 value, and the "distance" of their current test power from the theoretical power envelope.
- **Action space (A):** The pedagogical interventions available, such as: Level 1 (visual cue) — highlighting the shift in R_c^* as θ_1 approaches θ^*; Level 2 (Socratic hint) — asking the student to compare $P(X = x \mid H_0)$ vs. $P(X = x \mid H_1)$ for specific tail values; Level 3 (scaffolded problem) — presenting a simplified Binomial example to re-establish the likelihood ratio logic before returning to the Cauchy case.
- **Reward function (R):** The system receives a positive reward when the student's subsequent parameter adjustments move their test performance closer to the point optimal efficiency.

Feedback triggering logic: The feedback is not merely rule-based; it is triggered by a threshold-policy. When the engagement data logs a "Stagnation Event" — defined as three consecutive simulations where the student fails to recognize the change in rejection region shape at $\theta^* \approx 12.63$ — the RL engine overrides the standard display to provide a targeted visual explanation of the quadratic inequality. This technical depth ensures that the machine learning algorithm is operationalized to address specific misconceptions, such as the "Symmetry Bias" often seen in small-sample econometrics.

3. Significance Tests

We consider the problem of hypothesis testing in the simplest context. For example, we flip a coin N times and observe the outcomes. We think that the coin is fair but suspect that it might not be. We would like to test the null hypothesis that the coin is fair, with H (heads) and T (tails) each having probability 50%. Let X be the random variable which counts the number of heads within a sequence of N trials. Then $X \sim \text{Bi}(N, p)$, i.e. X has a Binomial distribution with parameters N and p, where N is the number of trials and p is the probability of heads. Outcomes which are likely under the null favour the null hypothesis, while unlikely outcomes cast doubt on it. For instance, consider Table 3.1 displaying the probabilities for a $\text{Bi}(10, 0.5)$ random variable.

h	$P(X = h)$	h	$P(X = h)$
0	0.001	6	0.205
1	0.010	7	0.117
2	0.044	8	0.044
3	0.117	9	0.010
4	0.205	10	0.001
5	0.246		

Table 3.1: Probabilities for a $\text{Bi}(10, 0.5)$

The central values 3, 4, 5, 6, 7 are likely and hence have higher probabilities, while the values far from the centre of 5 (tail values) — 0, 1, 2 and 8, 9, 10 — have low probabilities. The function $l(x, \theta)$ which gives the probability of the outcome x when the parameter is θ is called the likelihood function. Table 3.1 is a table of values of the likelihood function with parameter $p = 0.5$.

In significance testing, we have to decide upon a value c such that all likelihoods above c lead to acceptance of the null hypothesis, while all likelihoods below c lead to rejection. If we set $c = 0.02$ (or 2%), then outcomes $X = 0, 1$ and $X = 9, 10$ have probabilities less than c and

lead to rejection of the null hypothesis. In formal terms, $R = \{0,1,9,10\}$ is the rejection region and $A = \{2,3,4,5,6,7,8\}$ is the acceptance region.

The question is: how should we choose the constant c? This constant determines the probability of Type I error. A Type I error occurs when we reject a null hypothesis that is true. It is easy to see that the probability of Type I error is the probability of the rejection region under the null hypothesis. For the example given above, due to symmetry, it is $0.001 + 0.010 + 0.010 + 0.001 = 0.022$ or 2.2%. This quantity is also called the p-value of the test.

In the classical theory of hypothesis testing, the probability of a Type I error is also called the "significance level" of a test. The constant c is chosen to achieve conventional significance levels such as 5% or 1%. For example, the constant $c = 0.02$ leads to a p-value of 2.2%, which achieves the conventional significance level of 5% but not 1%. If we want to achieve a significance level of 1% then we must set $c = 0.005$ (i.e. 0.5%). In that case, the rejection region contains only $\{0,10\}$ and the p-value will be 0.002, i.e. 2 in 1000, which is less than 1%.

4. The Likelihood Ratio (LR)

Obviously, we would like to minimize the probability of Type I error. This could easily be done by never rejecting the null hypothesis. If the rejection region is the empty set, then we will never reject the null hypothesis and the probability of Type I error will be 0%.

In order to see that this is not a good test, we have to think seriously about alternatives to the null hypothesis. A Type II error occurs when the null hypothesis is accepted when it is false. For pedagogical purposes, in this paper the case of a simple null versus a simple alternative hypothesis is considered. A "simple" hypothesis is one with only one possible distribution.

Consider the random variable X which counts the number of heads in N trials, $X \sim \text{Bi}(N,p)$. The null hypothesis that the coin is fair is $H_0 : p = 0.5$. Suppose the alternative is $H_1 : p = 0.6$. Now, how should we construct the rejection region? The points $\{0,1\}$ which were in the rejection region for the symmetric test should not be used to reject H_0 here, since even though these values are unlikely under the null, they are even more unlikely under the alternative: $P(X = 0 \mid H_0;\ p = 0.5) = 0.000977$, while $P(X = 0 \mid H_1;\ p = 0.6) = 0.000105$. Getting 0 is ten times more likely under the null than under the alternative. When there is a specific alternative, we need to compare the probability of the outcome under the null and the alternative. This is called the likelihood ratio — the ratio of the probability under the alternative to the probability under the null. This is tabulated in Table 4.1.

h	$P(X{=}h \mid p{=}0.5)$	$P(X{=}h \mid p{=}0.6)$	Ratio	h	$P(X{=}h \mid p{=}0.5)$	$P(X{=}h \mid p{=}0.6)$	Ratio
0	0.000977	0.000105	0.107	6	0.205078	0.250823	1.223
1	0.009766	0.001573	0.161	7	0.117188	0.214991	1.835
2	0.043945	0.010617	0.242	8	0.043945	0.120932	2.752
3	0.117188	0.042467	0.362	9	0.009766	0.040311	4.128
4	0.205078	0.111477	0.544	10	0.000977	0.006047	6.192
5	0.246094	0.200658	0.815				

Table 4.1: The likelihood ratio for $\text{Bi}(10,p)$

The last column is also called the "odds ratio". The odds ratio for $X = 0$ is 0.107 (10.7%), meaning the likelihood of observing $X = 0$ is only 10.7% under the alternative relative to the null. In contrast, the outcome $X = 10$ is about 6 times as likely under the alternative as under the null.

An intuitively plausible principle for hypothesis testing is to reject the null hypothesis whenever the likelihood ratio is above c and accept it whenever the ratio is below c. Suppose we set $c = 1$. Then we reject for $\{6,7,8,9,10\}$ and accept otherwise. If the null is true, the probability of rejection is 37.7%. The probability of a Type II error (accepting the null when false) is 36.7%. The *power* of the test is $1 - 36.7\% = 63.3\%$.

As c increases, the probability of Type I error is reduced but that of Type II error is increased. The trade-off between these two probabilities is depicted in Table 4.2. If we make $c = 0$ and reject for all observations, then Type I error probability is 100% and Type II error probability is 0. As we increase c, we decrease Type I error probabilities and increase Type II errors until we reach 0% for Type I together with 100% for Type II error. Because the observation X is discrete, the trade-off curve has jumps in it. For continuous random variables, this curve is

Type I	0	0.001	0.011	0.055	0.172	0.377	0.623	0.828	0.945	0.989	0.999	1
Type II	1	0.994	0.954	0.833	0.618	0.367	0.166	0.055	0.012	0.002	0	0

Table 4.2: Type I and Type II error probabilities

smooth and concave, lying underneath the diagonal from $(0,1)$ to $(1,0)$ in the unit square. The Neyman–Pearson Lemma [7] says that tests based on the likelihood ratio are the best possible.

4.1. Neyman–Pearson (NP) test

Suppose that observation $X = x$ has likelihood $l(x,\theta)$ where $\theta = \theta_0$ under the null hypothesis and $\theta = \theta_1$ under the alternative hypothesis. Then the likelihood ratio (LR) is

$$LR(x,\theta_1,\theta_0) = \frac{l(x,\theta_0)}{l(x,\theta_1)}.$$

It is logical to reject the null hypothesis for large values of this ratio and to accept it for small values. For each constant c, the likelihood ratio test defines the rejection region $R^*(c) = \{x : LR(x, \theta_0, \theta_1) \geq c\}$.

Let $\alpha = P\big(x \in R^*(c) \mid \theta = \theta_0\big)$ be the probability of Type I error. Furthermore, let R be any other rejection region with $P(x \in R \mid \theta = \theta_0) \leq \alpha$. Then according to the Neyman–Pearson Lemma [7], R must have higher Type II error probability: $P(x \notin R \mid \theta = \theta_1) \geq P\big(x \notin R^*(c) \mid \theta = \theta_1\big)$. This likelihood ratio test is also known as the Neyman–Pearson (NP) test.

No test can improve on the NP test in terms of both Type I and Type II error probabilities. Since minimizing Type II error is equivalent to maximizing power, the Neyman–Pearson Lemma gives the most powerful test of significance level α for any fixed alternative θ_1.

4.2. Uniformly most powerful test (UMP)

Although the NP test is the most powerful test at alternative θ_1, its form and the value of constant c depend on the value of the parameter in the alternative hypothesis. If a single test can achieve the highest possible power against *all* alternatives, then this test is known as the uniformly most powerful test (UMP). It occurs when the NP test statistic and constant c are the same for all alternatives θ_1.

4.3. Locally most powerful test (LMP)

LMP tests are NP tests which maximize power very close to the null hypothesis, i.e. an NP test which optimizes power as $\theta \to \theta_0$. The choice of LMP as an alternative to UMP had been very popular [6, 8]. The revolution in computer technology has removed the old requirement of a test statistic to have a known asymptotic distribution, enabling researchers to calculate critical values of many tests using simulations.

4.4. Berenblut–Webb type test (BW)

Berenblut and Webb [31] recommended using the NP test when taking the limit of θ to the farthest point from the null hypothesis. BW is the complete opposite of the LMP test, since it maximizes power at the farthest point from the null.

4.5. Fraser et al. and King's test

The natural intermediate value of the alternative hypothesis was suggested by Fraser et al. [32] and King [33] after considering the subpar performance of both the LMP and BW tests. The underlying idea — specifically, that intermediate values have better performance than either extreme — is correct. However, the selection of appropriate intermediate values depends on the sample size.

4.6. Point optimal test

Consider an observation $X = x$ which has likelihood $l(x, \theta)$, where $\theta = \theta_0$ under the null hypothesis. For any alternative $\theta = \theta_1$, the Neyman–Pearson Lemma [7] shows how to compute the most powerful test of level α. The power of this test, $\pi_\alpha(\theta_1)$, is the maximum power that any test with significance level α can have. This function is called the power envelope and serves as the standard of comparison for any test. The NP test is also called a Point Optimal (PO) test because it is the most powerful at a single point $\theta = \theta_1$ in the alternative space. It may not be very good at other points.

In this paper, a simple case where the uniformly most powerful (UMP) test does not exist is discussed, so the point optimal (PO) test must be computed. For this purpose, the problem of hypothesis testing of the Cauchy location parameter is considered.

An example of the Cauchy location parameter. To clarify the concept of the point optimal (PO) test, a single draw from a Cauchy distribution is used as a case study. Let X be a random variable with a Cauchy distribution with location parameter θ. The likelihood function for $X = x$ is:

$$l(x, \theta) = \frac{1}{\pi\big[1 + (x - \theta)^2\big]}, \qquad -\infty < x < \infty.$$

Consider testing $H_0 : \theta = 0$ against $H_A : \theta = \theta_1 > 0$. According to the Neyman–Pearson Lemma [7], the best test must reject the null hypothesis for all x for which:

$$LR(x, \theta_1, 0) = \frac{l(x, 0)}{l(x, \theta_1)} \leq c.$$

The constant c is chosen so that the probability of rejection under the null equals α. The rejection region is $R^*(c) = \{x : LR(x, \theta_1, 0) \leq c\}$. The defining inequality can be written as:

$$LR(x, \theta_1, 0) = \frac{1 + (x - \theta_1)^2}{1 + x^2} \leq c \implies 1 + (x - \theta_1)^2 \leq c(1 + x^2).$$

This can be re-written as the quadratic inequality in x:

$$(c - 1)x^2 + 2\theta_1 x + (c - 1 - \theta_1^2) \geq 0. \tag{4.1}$$

The rejection region has three possible shapes depending on c. If $c > 1$, then the quadratic is U-shaped and the rejection region lies outside the two roots. If $c = 1$, the inequality reduces to $x \geq \theta_1/2$. For $c < 1$, the quadratic has an inverted-U shape and the rejection region lies between the two roots of (4.1). These roots are:

$$r_1 = \frac{-\theta_1 - \sqrt{c\,\theta_1^2 - (c-1)^2}}{c-1} = \frac{-2(1+k^2) - k\sqrt{\theta_1^2+4}\,\theta_1}{k^2-\theta_1^2+4},$$
$$r_2 = \frac{-\theta_1 + \sqrt{c\,\theta_1^2 - (c-1)^2}}{c-1} = \frac{-2(1+k^2) + k\sqrt{\theta_1^2+4}\,\theta_1}{k^2-\theta_1^2+4}.$$

The detailed proof is in Appendix B.

5. The Rejection Region of the PO Test

5.1. The rejection region of the PO test of size α

Lemma 5.1. *For every α there exists a value θ^* such that the most powerful test of level α for $H_0 : \theta = 0$ versus $H_1 : \theta = \theta_1$ for all $\theta_1 > 0$ has the following form, where $k = \tan(\pi\alpha)$:*

- *If $\theta_1 < \theta^*$, then $r_2 < r_1$ and*

$$R^*(c) = \left\{ x : \frac{-2(1+k^2) + k\sqrt{\theta_1^2+4}\,\theta_1}{k^2-\theta_1^2+4} \leq x \leq \frac{-2(1+k^2) - k\sqrt{\theta_1^2+4}\,\theta_1}{k^2-\theta_1^2+4} \right\}.$$

- *If $\theta_1 = \theta^*$, then $R^*(c) = \left\{ x : x \geq \dfrac{\theta^*}{2} \right\} = \left\{ x : x \geq \dfrac{1}{k} \right\}$.*
- *If $\theta_1 > \theta^*$, then $r_1 < r_2$ and*

$$R^*(c) = \left\{ x \leq \frac{-2(1+k^2) - k\sqrt{\theta_1^2+4}\,\theta_1}{k^2-\theta_1^2+4} \right\} \cup \left\{ x \geq \frac{-2(1+k^2) + k\sqrt{\theta_1^2+4}\,\theta_1}{k^2-\theta_1^2+4} \right\}.$$

An explicit formula for the critical threshold is $\theta^* = 2\tan\big(\tan(\pi\alpha)\big) = \dfrac{2}{k}$.

5.2. The rejection region of the PO test of size $\alpha = 0.05$

For $\alpha = 0.05$, we compute $\theta^* \approx 12.6275$, with $k = \tan(0.05\pi) \approx 0.1584$. The roots become:

$$r_1 = -2.025 - \frac{0.1584\sqrt{\theta_1^2+4}\,\theta_1}{-\theta_1^2 + 1.0125\theta_1 + 4},$$
$$r_2 = -2.025 + \frac{0.1584\sqrt{\theta_1^2+4}\,\theta_1}{-\theta_1^2 + 1.0125\theta_1 + 4}.$$

Once the rejection regions are known, it becomes possible to calculate the power envelope. The threshold θ^* is not an arbitrary numerical artefact; it reflects a fundamental geometric transition in the likelihood ratio surface of the Cauchy distribution. Recall that $LR(x, \theta_1, 0) = [1 + (x-\theta_1)^2]/[1+x^2]$. For small positive $\theta_1 < \theta^*$, the bounded-interval rejection region $R^*(c) = \{x : r_2 \leq x \leq r_1\}$ captures the region of x-values that most strongly favour the alternative. As θ_1 increases toward θ^*, this interval stretches toward the right tail. At exactly $\theta_1 = \theta^*$, the most powerful test collapses to a one-sided ray $R^*(c) = \{x : x \geq \theta^*/2\}$. Beyond θ^*, the fat tails of the Cauchy distribution play a decisive role: observations in the left tail are also informative against the null, and the rejection region splits into two disjoint rays, $R^*(c) = \{x \leq r_1\} \cup \{x \geq r_2\}$. Concretely, for $\alpha = 0.05$, $k = \tan(0.05\pi) \approx 0.1584$ yields $\theta^* = 2/k \approx 12.6275$.

6. The Power Envelope

6.1. The power envelope (general)

Theorem 6.1. *Let $\pi_\alpha(\theta_1)$ be the maximum power attainable at $\theta_1 > 0$ by any test of level α for $H_0 : \theta = 0$ vs. $H_1 : \theta = \theta_1$. With $k = \tan(\pi\alpha)$ and $\theta^* = 2/k$:*

$$\pi_\alpha(\theta_1) = \begin{cases} \dfrac{1}{\pi}\left[\arctan\left(\dfrac{-2(1+k^2)-k\sqrt{\theta_1^2+4}\,\theta_1}{k^2-\theta_1^2+4}-\theta_1\right)-\arctan\left(\dfrac{-2(1+k^2)+k\sqrt{\theta_1^2+4}\,\theta_1}{k^2-\theta_1^2+4}-\theta_1\right)\right] & \text{if } \theta_1<\theta^*, \\ \dfrac{1}{2}-\dfrac{1}{\pi}\arctan\left(\dfrac{\theta_1}{2}\right) & \text{if } \theta_1=\theta^*, \\ 1-\dfrac{1}{\pi}\left[\arctan\left(\dfrac{-2(1+k^2)+k\sqrt{\theta_1^2+4}\,\theta_1}{k^2-\theta_1^2+4}-\theta_1\right)-\arctan\left(\dfrac{-2(1+k^2)-k\sqrt{\theta_1^2+4}\,\theta_1}{k^2-\theta_1^2+4}-\theta_1\right)\right] & \text{if } \theta_1>\theta^*. \end{cases} \tag{6.1}$$

6.2. The power envelope at size $\alpha = 0.05$

We specialize (6.1) to $\alpha = 0.05$, giving $\theta^* \approx 12.6275$. The power envelope $\pi_{0.05}(\theta_1)$ is:

$$\pi_{0.05}(\theta_1) = \begin{cases} \frac{1}{\pi}\arctan\left(\dfrac{\begin{array}{l}-0.3168\sqrt{\theta_1^2+4}\cdot(\theta_1^2+4)+0.6414\sqrt{\theta_1^2+4}\cdot\theta_1 \\ +(\theta_1^2+4)(2.0251\theta_1^4+10.1003\theta_1^2+8)-2.025\theta_1(\theta_1^2+3)\end{array}}{\theta_1^2+4}\right) & if\ \theta_1<12.6275, \\ [10pt]\frac{1}{2}-\frac{1}{\pi}\arctan\left(\frac{\theta_1}{2}\right) & if\ \theta_1=12.6275, \\ [10pt]\frac{1}{\pi}\left[-\arctan\left(\dfrac{\begin{array}{l}0.3168\sqrt{\theta_1^2+4}\cdot(\theta_1^2+4)+0.6414\sqrt{\theta_1^2+4}\cdot\theta_1 \\ +(\theta_1^2+4)(2.0251\theta_1^4+10.1003\theta_1^2+8)+2.025\theta_1(\theta_1^2+3)\end{array}}{\theta_1^2+4}\right)\right] & if\ \theta_1>12.6275. \end{cases}$$

Interpreting the power envelope in applied teaching. The power envelope $\pi_\alpha(\theta_1)$ represents the theoretical ceiling on what any test of level α can achieve at each alternative θ_1. In an applied teaching context, this distinction is critical because students often conflate the power envelope with the power of the test they happen to be studying. The interactive simulation in the LMS module is specifically designed to counter this misreading by overlaying the power envelope as a reference ceiling above the power curves of individual tests. A student who adjusts θ_1 and observes how the gap between their selected test's power and the envelope changes in real time develops an intuitive grasp of efficiency loss — something that static textbook figures cannot convey. Furthermore, without visualization, students frequently assume that power is a monotonically increasing function of the distance between null and alternative. The piecewise structure of $\pi_{0.05}(\theta_1)$ for the Cauchy, and the shift in regime at $\theta^* \approx 12.6275$, directly contradicts this assumption, illustrating that for heavy-tailed distributions optimal testing strategy must adapt qualitatively as the alternative moves further from the null.

7. Discussion and Implications for Educational Technology

The integration of rigorous econometric theory with AI-supported interactive simulations offers multiple benefits for teaching and learning in quantitative disciplines. First, by embedding the point optimal test and power envelope concepts into a dynamic, LMS-based environment, students can explore parameter variations in real time, making the learning experience more active and inquiry-driven. This approach supports constructivist learning principles, where students build understanding through direct interaction with mathematical models rather than passively consuming information.

Second, the adaptive feedback system powered by AI analytics transforms the simulation into more than just a visualization tool; it becomes a personalized tutor. By analysing patterns of student engagement, error frequency, and parameter exploration sequences, the system can identify specific conceptual bottlenecks and deliver tailored explanations or practice opportunities. Such adaptive systems have been shown to improve learning efficiency in mathematics and statistics education [13, 18] and the implementation demonstrates their feasibility in advanced econometrics contexts.

Third, from an instructional perspective, the data generated by the system provides valuable insights for educators. Instructors can monitor engagement metrics, identify students at risk of falling behind, and adapt course delivery in response to emerging learning patterns as displayed in Appendix A and Appendix B. This aligns with the growing emphasis on learning analytics in higher education, where instructional decisions are increasingly supported by data-driven evidence.

Finally, the broader implication of this study is that highly abstract mathematical topics can be made approachable through thoughtfully designed educational technology. While the current implementation focuses on the Cauchy distribution as a tractable case, the framework can be extended to a range of statistical models and hypothesis-testing scenarios, thereby broadening its applicability across econometrics, statistics, and other quantitative disciplines.

7.1. Future work

Building on the present implementation, future work will focus on expanding the AI capabilities of the simulation platform to further enhance personalization and learner engagement. One direction is the integration of predictive learning analytics, where machine learning models forecast student performance trajectories based on early interaction data. Another avenue is the incorporation of natural language

processing (NLP) to provide more conversational, context-aware explanations of statistical concepts, mirroring the behaviour of an intelligent tutoring system. Additionally, gamification elements, such as achievement badges and leaderboards, may increase student motivation and sustained engagement. From a scalability perspective, the framework could be adapted to other hypothesis-testing scenarios and integrated into MOOCs or open-access learning platforms.

Article Information

Acknowledgements: The authors would like to express their sincere thanks to the editor and the anonymous reviewers for their helpful comments and suggestions.

Author Contributions: Conceptualization: AUIK, MAB, AY; Methodology: MAB; Software: AUIK, MAB; Validation: MAB, AY; Formal analysis: MAB, AY; Investigation: AUIK, MAB; Resources: AUIK, MAB, AY; Data curation: AUIK, MAB, AY; Writing – original draft: AUIK, MAB, AY; Writing – review & editing: AUIK, MAB, AY; Visualization: AUIK, MAB; Project administration: AUIK.

Artificial Intelligence Statement: No artificial intelligence has been utilized in the writing of any portion of the paper. Grammarly V2 was used for language editing.

Conflict of Interest Disclosure: No potential conflict of interest was declared by authors.

Plagiarism Statement: This article was scanned by the plagiarism program.

References

[1] J. Durbin, G. S. Watson, *Testing for serial correlation in least squares regression: I*, Biometrika, **37**(3/4) (1950), 409–428. https://doi.org/10.2307/2332391
[2] J. Durbin, G. S. Watson, *Testing for serial correlation in least squares regression: II*, Biometrika, **38**(1/2) (1951), 159–178. https://doi.org/10.1093/biomet/38.1-2.159
[3] T. U. Islam, *Ranking of normality tests: An appraisal through skewed alternative space*, Symmetry, **11**(7) (2019), 872. https://doi.org/10.3390/sym11070872
[4] A. U. I. Khan, W. M. Khan, M. Hussan, *Most stringent test of null of cointegration: A Monte Carlo comparison*, Commun. Stat. Simul. Comput., **51**(4) (2022), 2020–2038. https://doi.org/10.1080/03610918.2019.1691229
[5] W. M. Khan, A. U. I. Khan, *Most stringent test of independence for time series*, Commun. Stat. Simul. Comput., **49**(11) (2020), 2808–2826. https://doi.org/10.1080/03610918.2018.1527350
[6] A. Zaman, *Statistical Foundations for Econometric Techniques*, PIDE, 1996. https://www.academia.edu/1870016/Statistical_Foundations_for_Econometric_Techniques
[7] J. Neyman, E. S. Pearson, *IX. On the problem of the most efficient tests of statistical hypotheses*, Philos. Trans. R. Soc. Lond. Ser. A, **231** (1933), 289–337. https://doi.org/10.1098/rsta.1933.0009
[8] E. L. Lehmann, J. P. Romano, *Testing Statistical Hypotheses*, Springer, New York, 2005.
[9] T. U. Islam, *Stringency-based ranking of normality tests*, Commun. Stat. Simul. Comput., **46**(1) (2017), 655–668. https://doi.org/10.1080/03610918.2014.977916
[10] A. U. I. Khan, A. Zaman, *Theoretical and Empirical Comparisons of Cointegration Tests*, Ph.D. dissertation, International Islamic University, 2017.
[11] G. Casella, R. Berger, *Statistical Inference*, CRC Press, 2024. https://doi.org/10.1201/9781003456285
[12] D. M. Lane, S. C. Peres, *Interactive simulations in the teaching of statistics: Promise and pitfalls*, in Proceedings of the Seventh International Conference on Teaching Statistics, International Statistical Institute, Voorburg, 2004.
[13] R. S. Baker, P. S. Inventado, *Educational data mining and learning analytics*, in Springer eBooks, Springer, 2014, 61–75. https://doi.org/10.1007/978-1-4614-3305-7_4
[14] J. T. McClave, P. G. Benson, T. Sincich, *Statistics for Business and Economics*, Pearson Education, 2008.
[15] J. Garfield, D. Ben-Zvi, *Developing Students' Statistical Reasoning: Connecting Research and Teaching Practice*, Springer, 2008.
[16] K. Kuosa, D. Distante, A. Tervakari, et al., *Interactive visualization tools to improve learning and teaching in online learning environments*, Int. J. Distance Educ. Technol., **14**(1) (2016), 1–21. https://doi.org/10.4018/ijdet.2016010101
[17] R. Nkambou, J. Bourdeau, R. Mizoguchi, *Introduction: What are intelligent tutoring systems, and why this book?* in Studies in Computational Intelligence, Springer, 2010, 1–12. https://doi.org/10.1007/978-3-642-14363-2_1
[18] M. Chi, K. VanLehn, D. Litman, et al., *An evaluation of pedagogical tutorial tactics for a natural language tutoring system: A reinforcement learning approach*, Int. J. Artif. Intell. Educ., **21**(1–2) (2011), 83–113. https://doi.org/10.3233/jai-2011-014
[19] M. C. Murray, *Bridging the generative AI literacy gap: A guide to introducing prompt engineering in university courses*, Issues Informing Sci. Inf. Technol., **22** (2025), Article ID 10.
[20] C. Barreto, *Making a preliminary case for a universal course on AI literacy: An overview*, in InSITE 2025, Informing Science + IT Education Conferences, Hiroshima, **18** (2025), Article ID 18.
[21] P. A. Rospigliosi, *Teaching in the second machine age*, Interact. Learn. Environ., **24**(8) (2016), 1741–1743. https://doi.org/10.1080/10494820.2017.1262142
[22] T. Karakose, T. Tülübas, *How can ChatGPT facilitate teaching and learning: Implications for contemporary education*, Educ. Process Int. J., **12**(4) (2023), 7–16. https://doi.org/10.22521/edupij.2023.124.1
[23] A. Duran, U. F. Ermiş, *Integrating generative AI in education: Implications for school leadership from policy documents through Bolman and Deal's four frame theory*, Bartın Univ. J. Fac. Educ., **14**(2) (2025), 416–440.
[24] P. A. Rospigliosi, *What is the role of ChatGPT and other large language model AI in higher education?* Interact. Learn. Environ., **32**(2) (2024), 393–394. https://doi.org/10.1080/10494820.2024.2330836
[25] D. Soto, M. Higashida, S. Shirai, et al., *Enhancing learning dynamics: Integrating interactive learning environments and ChatGPT for computer networking lessons*, Procedia Comput. Sci., **246** (2024), 3595–3604.
[26] A. Durgungoz, A. Kharrufa, *ChatGPT is like a study buddy, a teacher and sometimes just a friend: A longitudinal exploration of students' interactions, perception and acceptance*, Interact. Learn. Environ., **34**(2) (2026), 758-777. https://doi.org/10.1080/10494820.2025.2509276
[27] X. Y. Wu, J. D. Radloff, I. H. Yeter, et al., *Designing artificial intelligence chatbots for self-regulated learning from a systematic review based on Habermas's three interests*, Interact. Learn. Environ., (2025), (Early Access). https://doi.org/10.1080/10494820.2025.2563086
[28] A. Strzelecki, *To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology*, Interact. Learn. Environ., **32**(9) (2024), 5142–5155. https://doi.org/10.1080/10494820.2023.2209881
[29] S. Bangay, S. McKenzie, G. Wood-Bradley, *Teaching-focused rapid authoring of educational extended reality (eduXR) experiences*, Interact. Learn. Environ., **34**(3) (2026), 1626-1646. https://doi.org/10.1080/10494820.2025.2527166
[30] I. du Preez, L. Jacobs, *Pedagogical and visual design principles for online learning material: Insights from quality guiding documents in diverse educational contexts*, Interact. Learn. Environ., **34**(3) (2026), 1435-1451. https://doi.org/10.1080/10494820.2025.2523390
[31] I. Berenblut, G. I. Webb, *A new test for autocorrelated errors in the linear regression model*, J. R. Stat. Soc. Ser. B, **35**(1) (1973), 33–50. https://doi.org/10.1111/j.2517-6161.1973.tb00932.x

[32] D. A. S. Fraser, I. Guttman, G. P. H. Styan, *Serial correlation and distributions on the sphere*, Commun. Stat. Theory Methods, **5**(2) (1976), 97–118. https://doi.org/10.1080/03610927608827336
[33] M. L. King, *Towards a Theory of Point Optimal Testing*, AgEcon Search, 1988. https://doi.org/10.22004/ag.econ.266861

Appendix

A. SWOT analysis of the simulation module

The SWOT analysis presented in Appendix A provides a roadmap for the ongoing refinement of the AI-supported simulation module. While the current "Interactive learning experience" fosters improved understanding through active and visual learning, the identified "Limited scope" — specifically the focus on the Cauchy distribution — indicates a need to broaden statistical concept coverage in future iterations. Furthermore, addressing "Student resistance" and potential over-reliance on simulation will require instructional strategies that mitigate dependence on AI feedback and prevent the misinterpretation of results. These insights underscore that the educational significance of the tool depends not just on its technological presence, but on how it is positioned within the pedagogical structure.

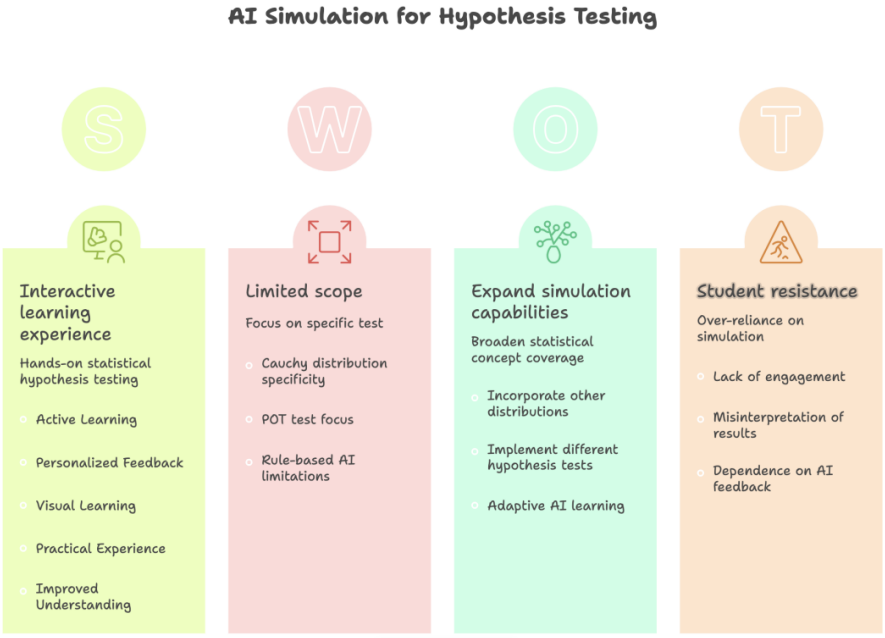

Figure A.1: LMS Dashboard — Overview of the AI-supported simulation module integrated into Canvas LMS, showing the interactive hypothesis testing environment.

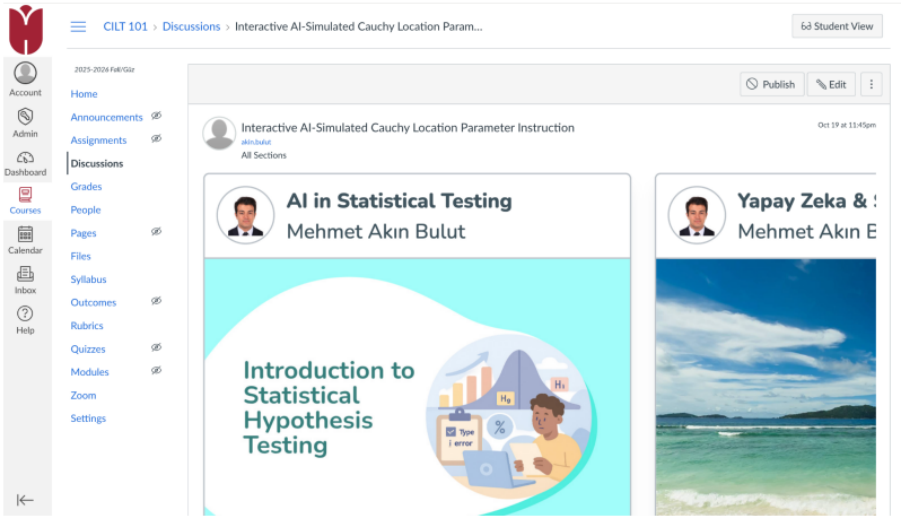

Figure A.2: LMS Dashboard — Real-time power envelope visualization and rejection region display as students adjust the alternative hypothesis parameter θ_1.

Figure A.3: LMS Dashboard — Instructor heatmap of misconceptions across the cohort, identifying students struggling at the θ^* phase transition.

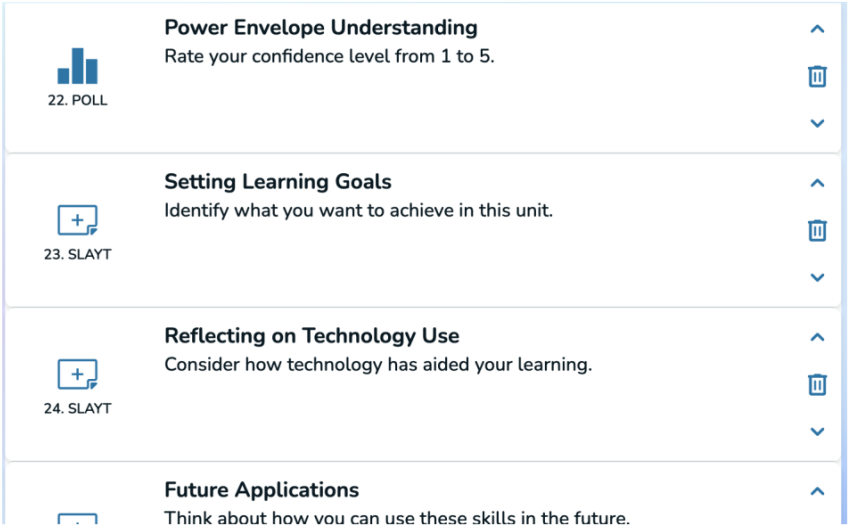

Figure A.4: Adaptive feedback interface — RL-based hint delivery triggered by a "Stagnation Event" in student interaction telemetry.

B. Proof of the roots r_1 and r_2

Let $X = x$ be a single draw from the Cauchy distribution. The density function with location parameter θ is $f(x) = \{\pi[1+(x-\theta)^2]\}^{-1}$. The likelihood functions under the null and alternative hypotheses are:

$$l(x,0) = \frac{1}{\pi(1+x^2)}, \qquad l(x,\theta_1) = \frac{1}{\pi[1+(x-\theta_1)^2]}.$$

The likelihood ratio test statistic (NP test) is:

$$LR(x,\theta_1) = L(x,0,\theta_1) = \frac{LH(x,0)}{LH(x,\theta_1)} = \frac{1+(x-\theta_1)^2}{1+x^2}.$$

At fixed size α there exists a constant c such that:

$$P\left(\frac{1+(x-\theta_1)^2}{1+x^2} \leq c\right) = \alpha. \tag{B.1}$$

Considering only the inequality in (B.1): $(c-1)x^2 + 2\theta_1 x + (c-1-\theta_1^2) \geq 0$, so that

$$P\big((c-1)x^2 + 2\theta_1 x + (c-1-\theta_1^2) \geq 0\big) = \alpha. \tag{B.2}$$

Considering only the inequality in (B.2):

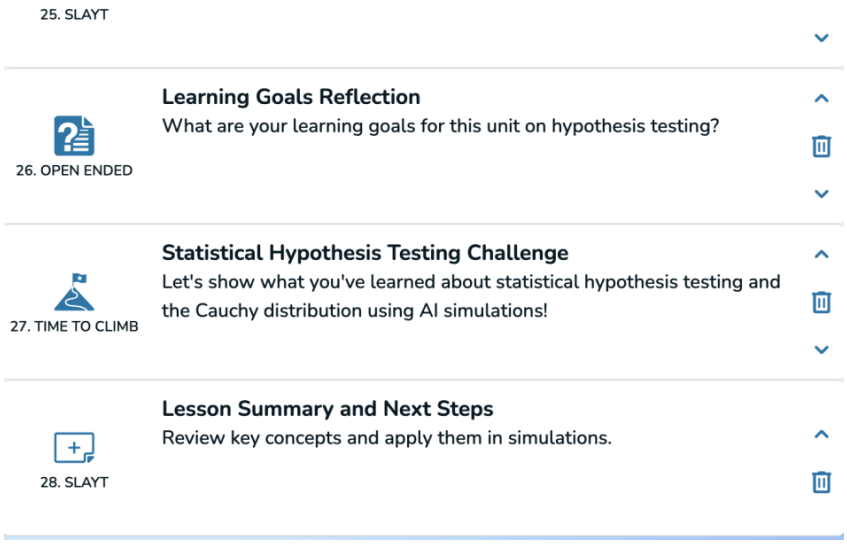

Figure A.5: Simulation module — Point-Probe feature showing the exact likelihood ratio $L(x,\theta_1,0)$ for a selected observation x, helping students visualize heavy-tail behaviour of the Cauchy distribution.

Case 1: $c = 1 \Rightarrow x \geq \theta_1/2$. The rejection region is $R*(c) = \{x : x \geq \theta_1/2\}$. By definition,

$$\alpha = \int_{\theta_1/2}^{\infty} \frac{dx}{\pi(1+x^2)} = \frac{1}{\pi}\left[\frac{\pi}{2} - \arctan\left(\frac{\theta_1}{2}\right)\right],$$

which gives $\theta_1 = \frac{2}{\tan(\pi\alpha)} = \theta^*$.

For Cases 2 and 3, the quadratic inequality in (B.2) yields two roots r_1 and r_2:

$$r_1 = \frac{-\theta_1 - \sqrt{c\theta_1^2 - (c-1)^2}}{c-1}, \tag{B.3}$$

$$r_2 = \frac{-\theta_1 + \sqrt{c\theta_1^2 - (c-1)^2}}{c-1}. \tag{B.4}$$

Case 2: $c > 1 \Rightarrow \theta_1 > \theta^*$ and $r_1 < r_2$; rejection region $R^*(c) = \{x \leq r_1\} \cup \{x \geq r_2\}$.

Case 3: $c < 1 \Rightarrow \theta_1 < \theta^*$ and $r_1 > r_2$; rejection region $R^*(c) = \{x : r_2 \leq x \leq r_1\}$.

Computing c. Treating r_1 and r_2 as quadratic roots of (B.2):

$$r_1 + r_2 = \frac{-2\theta_1}{c-1}, \qquad r_1 r_2 = \frac{c-1-\theta_1^2}{c-1}. \tag{B.5}$$

For $c < 1$ (Case 3):

$$\alpha = \int_{r_2}^{r_1} \frac{dx}{\pi(1+x^2)}.$$

Using the arctangent subtraction formula with $k = \tan(\pi\alpha)$:

$$k = \frac{r_1 - r_2}{1 + r_1 r_2} \quad \Rightarrow \quad (r_1 - r_2)^2 = k^2(1 + r_1 r_2)^2. \tag{B.6}$$

Since $(r_1 - r_2)^2 = (r_1 + r_2)^2 - 4r_1 r_2$, using (B.5):

$$(r_1 - r_2)^2 = \frac{4c\theta_1^2 - (c-1)^2}{(c-1)^2}. \tag{B.7}$$

Substituting (B.5) and (B.7) into (B.6) and simplifying:

$$-4(1+k^2)(c-1)^2 + 4\theta_1^2(1+k^2)(c-1) + \theta_1^2(4 - k^2\theta_1^2) = 0.$$

Solving for c:

$$c = 1 + \frac{\theta_1^2}{2} \pm \frac{\theta_1}{2}\sqrt{\theta_1^2 + 4(1+k^2)}.$$

Since we considered $c < 1$, we take:

$$c = 1 + \frac{\theta_1^2}{2} - \frac{\theta_1}{2}\sqrt{\theta_1^2 + 4(1+k^2)}.$$

Substituting this value of c into (B.3) and (B.4) finally yields:

$$r_1 = \frac{-2(1+k^2) - k\sqrt{\theta_1^2+4}\,\theta_1}{k^2 - \theta_1^2 + 4},$$

$$r_2 = \frac{-2(1+k^2) + k\sqrt{\theta_1^2+4}\,\theta_1}{k^2 - \theta_1^2 + 4}.$$