

Evaluation of Scientific Evidence Based on Likelihood Ratio as a Function of Measurement Uncertainty for Different Continuous Distribution Functions

Israa Abdulameer Resen*

*Department of Mobile Communications and Computing Engineering, College of Engineering,
University of Information Technology and Communications (UoITC), Baghdad, Iraq*

Abstract For the calculation of the likelihood ratio (LR) in cases of continuous forensic evidence, one must decide whether to use exact integration across the instrument uncertainty or simplified PDF approximations. In this paper, we provide theoretical bounds and empirical support for the equivalence of exact and approximate LR calculations. It is shown that in cases of normally distributed evidence, the two approaches provide almost equivalent LR values if the relative instrumental uncertainty is less than 2% of the measured value. For the cases of small sample sizes ($N = 3, DOF = 2$) and modelled by Student's t-distribution, the convergence point becomes 4.5%. Validation of the approach based on the empirical data set provided from a real case of glass refractive index (RI) measurement using $GRIM3$ instrumentation ($ASTME1967$ standards) is conducted. The data show that the relative uncertainty for this operation equals 0.0033% and is comfortably operating inside the bounds, providing indistinguishable LR values (" LR " = 7.14 *vs.* 7.15). Finally, we address the problem of 'rare type', showing that it is resolved by Student's t-distributions that do not allow unnecessary inflation of LR at tail events.

Keywords Bayesian Statistics; Likelihood Ratio; Error of Measurement; Shape of Continuous Distribution; Forensic Data

DOI: 10.19139/soic-2310-5070-4425

1. Introduction

The quantification of the weight of scientific evidence has become an integral part of modern-day forensic science in which objective statistical and probabilistic reasoning becomes increasingly preferable to subjective expertise. Among the existing probabilistic frameworks for evidence assessment, the LR has become the internationally standardised way of expressing the weight of the forensic evidence in the Bayesian framework ([2]; [23]). In contrast to providing an assertion regarding whether the suspect is either guilty or innocent, the likelihood ratio provides an estimate of how many times more likely the evidence is under the alternative hypothesis compared to the null one by enabling a rational combination of the scientific evidence and specific case-related facts ([3]; [10]). As a result, the LR approach has become widely chosen by forensic disciplines such as DNA profiling, glass evidence, fibre examination, trace evidence, fingerprints, footprints, digital forensics and forensic engineering ([7]; [14]; G. S. [13]; [14]; [6]; [18]) resulting in different kinds of continuous scientific evidence owing to the mathematical consistency and interpretative transparency of the LR methodology ([22]; [20]). The latest advancements have strengthened the key role of Bayesian logic and likelihood ratios in forensic reasoning while also pointing out the importance of using adequate statistical models and proper accounting for the measurement uncertainty and sample size in the process of evaluating the evidence ([4]; [22]). Nevertheless, an important

*Correspondence to: Israa Abdulameer Resen (Email: israa.resen@uoitc.edu.iq). Department of Mobile Communications and Computing Engineering, College of Engineering, University of Information Technology and Communications (UoITC), Baghdad, Iraq.

methodological problem has not been thoroughly investigated. Namely, while the current research has mostly focused on constructing, validating and interpreting the likelihood ratios for different kinds of forensic evidence, relatively little effort has been made to investigate the effect of the measurement uncertainty and sample size on the numerical properties of the likelihood ratio itself under different continuous probability distributions. In particular, no detailed guidelines regarding the realistic conditions under which the approximation of the frequency distribution can be used instead of the precise calculation of the probability for the normal and Student's *t*-distributions have been provided. Such a shortcoming is especially relevant in forensic applications, including small sample sizes, since the choice of the underlying probability distribution and measurement uncertainty may significantly affect the reported evidential value.

In the case of more recent works, attention has been placed on enhancing the *LR* calculations via probabilistic modelling, score methods, calibration procedures, and empirically validated approaches to ensure the accurate interpretation of the results ([16]; [9]). These studies have enhanced the likelihood ratio systems to a great extent both theoretically and in practical terms. Nevertheless, their importance lies mostly in the development of models, calibration, and validation instead of the study of the mathematical impact of measurement uncertainties and sample sizes on the *LR* calculations of continuous evidence. In practice, continuous forensic measurements are normally assumed to follow the normal distribution. However, in cases where there is not enough data to make such assumptions, it is more appropriate to use the Student's *t*-distribution. Also, only limited attention has been paid to the comparison of the results obtained by using probability integration and approximation of the probability density function for various measurement uncertainties and various probability distributions of different sample sizes.

Recent studies have shown that the reliability of the likelihood ratio assessment is influenced by both the choice of the statistical model used and the proper consideration of the measurement uncertainties. Morrison ([13]) stressed the role of empirical verification as necessary to make sure that *LR* systems produce valid and reliable information. In turn, [9] and [16] showed that the *LR* systems must be verified by means of calibration and validation prior to applying them in the practice of casework. The recommendations of such international organisations as *ASTM*, *ENFSI* and *NIST* include the considerations of transparency, reproducibility and measurement uncertainties ([23]; [15]; [5]).

From a more general perspective on methodology, some statistical and mathematical models have shown the importance of performing uncertainty analysis, constructing a likelihood function, and estimating density within probabilistic systems. For example, in the case of interval arithmetic, it has been shown how to establish the parameter boundaries to deal with uncertainty in system evaluation ([21]). With regard to Bayesian methodology, the study of the choice of a prior has revealed its interdependency with the estimate ([12]). Moreover, recent methodologies comparing Bayes' theorem with the maximum likelihood approach have indicated the use of both methods for the estimation of continuous parameters under conditions of uncertain data ([8]).

In addition, past research efforts have been largely focused on the development and implementation of the theory and methodology of the likelihood ratio calculation. The formulation of practical criteria for choosing a specific approach to computation in relation to the conditions of uncertainty associated with the process of measuring forensic evidence is still unclear. In particular, there has been relatively little focus on the identification of conditions under which the approximation of the probability density function provides likelihood ratios that are statistically equal to the likelihood ratios computed through exact probability calculation. Hence, there is still a need for comprehensive research into the interaction of measurement uncertainty, sample size, probability function properties, and the behaviour of the likelihood ratio.

This review has illustrated that great progress has already been made regarding the use of *LR* approach in the assessment of forensic evidence. However, there are still some issues remaining to be resolved from the current literature review. First, the majority of literature works have concentrated on the discussion of the interpretive, validation and evaluation performance of likelihood ratio methods. Much less attention has been paid to the impact of measurement uncertainty on the value of likelihood ratio calculation in the case of continuous evidence. Second, despite the popularity of using normal and Student's *t*-distributions in statistics, few researchers have examined how the use of different probability distributions affects the convergence between exact probability integration and approximation through a probability density function based on the sample size. Third, no such practical analytical

framework has been proposed to determine the threshold uncertainty at which the approximation can replace the exact solution.

This study attempts to solve the above-stated problem of research and evaluates the impact of the uncertainty of evaluation on the LR of continuous evidence and the sample size. Instead of studying the theoretical background, validation and practical use of the LR structure, the work presents a mathematical comparison of the exact probability integration model and the approximation model based on the probability density function in the cases of the normal and Student's t -distribution. Special attention will be paid to the effect of decreasing maximum measurement error and the effect of different numbers of degrees of freedom on the convergence of the methods. Thus, it is expected to develop criteria of proficiency for recognising when it is possible to substitute the exact likelihood ratio calculation method for the approximate one without lowering its accuracy. As a result, the proposed statistical framework provides substantiated recommendations regarding the choice of the right way of the likelihood ratio evaluation depending on the specific features of the forensic evidence. In this way, the results obtained will enhance the efficiency of continuous forensic evidence evaluation, increase its statistical validity and allow applying Bayesian inference in forensic liability.

2. Method

Each piece of data related to a specific forensic issue can follow a special type of distributions like normal distribution, t -distribution, with population standard deviations known or unknown. The DOF , which are dependent on the sample size, are considered the main factor affecting the shape of the distribution function. The LR is calculated based on the ratio of two probabilities. Each of the two probabilities is estimated based on the area of a specific segment under the corresponding distribution function. This method is called the exact probability integration approach or simply the exact method. The corresponding distribution function must match the forensic data of the problem under consideration at the crime scene, which is dependent on the sample size provided as potential evidence. The LR can be generally defined in the form ([3]; [7]; [1]):

$$LR = \frac{P(E|H_1)}{P(E|H_2)} \quad (1)$$

Where “ E ” is the evidence event, H_1 represents the hypothesis supported by the prosecution, while H_2 represents the alternative hypothesis supported by the defence. The H_1 and H_2 are two mutually exclusive assumptions. $P(E|H_1)$ is defined as the probability of the evidence being at the crime scene, given H_1 is true (This is the probability of the evidence given the truth of the prosecution's case), while $P(E|H_2)$ represents the probability of the evidence being at the crime scene, given H_2 is true (This is the probability of the evidence given the truth of the defence case). If the LR is greater than unity $LR > 1$, this indicates that, based on the hypothesis of the numerator (guilt), greater weight is accorded than to the hypothesis of the denominator (innocence). On the other hand, if the ratio is less than unity $LR < 1$, this indicates that, based on the hypothesis of the denominator (innocence), greater weight is accorded than to the hypothesis of the numerator (guilt); it means that the proposition based on guilt carries less weight than that based on innocence. Finally, if the ratio is equal to unity $LR = 1$, at that point, the two hypotheses are on equal footing; that is, the evidence carries no weight in either direction (Adam, 2011).

To find the likelihood ratio, A statistical analysis must be conducted to determine the numerator $P(E|H_1)$ and the denominator $P(E|H_2)$ of the likelihood ratio for a given shape of distribution based on collected data related to the target evidence from the crime scene. Consider the following scenario: a set of measurements for any physical property, such as the glass refractive index in an armed robbery of a bank. Several glass fragments can be found at the crime scene due to the smashing of the bank security screen. The glass refractive index from these glass fragments is considered a control sample. The control sample yields a mean $\bar{X}$ and the standard deviation “ S ”. A single measurement “ x ” of the questioned material is also recorded, like a glass fragment on the defendant's coat, with a maximum error $\mp \Delta$ as uncertainty in the measurement (the uncertainty Δ is specified as the standard uncertainty in instrument measurements (u_{inst}) following calibrated analytical techniques (e.g., *ASTME1967*)). Here $P(E|H_1)$ will be less than unity and can be spoken of as the probability of finding the value “ $x \mp \Delta$ ”

from the probability distribution governing the control measurements. This can be written as $P(x \mp \Delta; \bar{X}, S)$ and $P(E/H_1) = P(x \mp \Delta; \bar{X}, S)$.

On the other side, the conditional probability $P(E/H_2)$ which is typically defined as the repetition of evidence within the environment. The physical meaning of $P(E/H_2)$ is the probability that the glass fragments found on the defendant's coat originated from another source as environmental contamination. Mathematically, calculating the probability of finding a material that matches specification "x" within a margin of uncertainty of $\mp \Delta$ in a target statistical population with a known distribution (characterised by a mean " μ " and a standard deviation " σ ") can be mathematically formulated as: $P(E/H_2) = P(x \mp \Delta; \mu, \sigma)$. As a result, the LR is therefore given by ([3]; [10]):

$$LR = \frac{P(E/H_1)}{P(E/H_2)} = \frac{P(x \mp \Delta; \bar{X}, S)}{P(x \mp \Delta; \mu, \sigma)} \quad (2)$$

At times, and in practice, the error of measurement $\mp \Delta$ is not known for several reasons, and only the single measurement "x" is known, like a single value of glass refractive index in a defendant coat without measurement error. In this case, some approximation is inevitable. As is known, the LR is based on the ratio of two probabilities, each representing the area of a specific segment under the corresponding target distribution function. This target distribution may follow the normal distribution or the t-distribution. As expected, the range of the maximum error $\mp \Delta$ in the measurement is small. It is shown that the probability, when calculated over a narrow interval, is approximately equal to the width of that interval multiplied by the value of the specific probability density function. This method is based on the rectangular strip approximation and is known as probability density function approximation or simply as an approximation method. The narrower the range, the more accurate this approximation becomes. If $f(x, \mu, \sigma)$ is representing the probability normal density distribution, as an example with " μ " and " σ ", and the range width is 2Δ , then the probability within this range $x \pm \Delta$ is:

$$P(x - \Delta \leq x \leq x + \Delta; \mu, \sigma) \approx f(x, \mu, \sigma) \times 2\Delta \quad (3)$$

The normal probability density distribution can be given as an exponential function ([19]):

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{(x - \mu)^2}{2\sigma^2} \right] \quad (4)$$

Then, the equation for the LR , where the maximum error " Δ " of measurement is sufficiently small, is given by:

$$LR = \frac{P(x = x \mp \Delta; \bar{X}, S)}{P(x = x \mp \Delta; \mu, \sigma)} \approx \frac{f(x, \bar{X}, S) \times 2\Delta}{f(x, \mu, \sigma) \times 2\Delta} \approx \frac{f(x, \bar{X}, S)}{f(x, \mu, \sigma)} \quad (5)$$

Equation 5 shows that the calculation is reduced to only finding the probability density functions of the target distribution shape at the point x , and the uncertainty in the measurement $\mp \Delta$ is not required in this case. It has been explicitly stated that 2Δ is cancelled by instrumental error since the instrumental error Δ is the same in H_1 and H_2 . A mathematical bound on the probability density approximation error can be set up using the second-order Taylor series expansion of a continuously differentiable density function $f(x)$ within the range $[x - \Delta, x + \Delta]$ of instrumental error.

$$\int_{x-\Delta}^{x+\Delta} f(t) dt = 2\Delta f(x) + \frac{\Delta^3}{3} f''(x) + O(\Delta^5) \quad (6)$$

Relative truncation error ϵ_{approx} for single density evaluation is thus limited as follows when using the rectangular approximation divided by $2\Delta f(x)$:

$$\epsilon_{approx} \approx \left| \frac{\Delta^2}{6} \frac{f''(x)}{f(x)} \right| \quad (7)$$

For a normally distributed background $N(\mu, \sigma^2)$, plugging in the function $f''(x)/f(x) = \frac{(x-\mu)^2 - \sigma^2}{\sigma^4}$ gives us a clear estimate of the relative likelihood error. If the relative measurement error Δ/x is less than 2%, or $\Delta \ll \sigma$,

then ϵ_{approx} approaches zero, thereby mathematically validating the fact that $f_1(x)/f_0(x)$ converges to the exact integral ratio without numerical loss of evidential integrity.

Equation 5 can be further simplified with the help of Eq. 4 to get.

$$LR \approx \frac{f(x, \bar{X}, S)}{f(x, \mu, \sigma)} \approx \frac{\frac{1}{S\sqrt{2\pi}} \exp - \frac{1}{2} \left(\frac{x - \bar{X}}{S} \right)^2}{\frac{1}{\sigma\sqrt{2\pi}} \exp - \frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2} = \frac{\sigma}{S} \exp \left(\frac{Z_\sigma^2 - Z_S^2}{2} \right) \quad (8)$$

Where, for convenience, we define $Z_S = \frac{x - \bar{X}}{S}$ and $Z_\sigma = \frac{x - \mu}{\sigma}$. If the distribution follows the t-shape, then the gamma function is used to follow the probability density distribution (Adam, 2011):

$$f(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi k} \Gamma\left(\frac{k}{2}\right) \left[\frac{x^2}{k} + 1 \right]^{\frac{(k+1)}{2}}} \quad (9)$$

Where $\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx$ and $p > 0$; here $k = N - 1 = DOF$ is the number of degrees of freedom. The mean and variance of the t-distribution are zero and $k/(k-2)$, respectively. The function of the t-distribution does not need to be dealt with in solving problems of the t-distribution because we only depend on the probability table or any available software. Equation 5 of a normal distribution can be applied in the same manner for the t-distribution, except that the degrees of freedom come into the picture. A mixed calculation may also be used where the numerator distribution may follow a t-distribution while the denominator distribution follows the normal curve. This mixed manipulation simulates the real practice in forensic science for the majority of problems. To take into account the uncertainty of estimating the small samples ($N = 3$), the predictive Location-Scale Student's t-distribution approach is used with $DOF = N - 1 = 2$. The use of Student's t-distribution rescaled with $S\sqrt{1 + 1/N}$ constitutes an intrinsic sensitivity analysis that explicitly takes into account sample parameter estimation uncertainty ($\bar{X}, S$)."

In this study, we try to compare the two methods given in Eq. 2 (Exact method) and Eq. 5 (Approximate Method) using two different continuous distributions, which are the normal distribution and the t-distribution. Further, the effect of the uncertainty value on the behaviour of Eq. 5 compared to Eq. 2 in these two different distributions. Minitab Software ([11]) is adopted in this research to find both the ordinate of the probability density function given in Eqs. 4 and 7 and the area under those curves. Numerical probability integrations have been performed in Minitab (version 20.4) with adaptive Gauss-Kronrod numerical quadrature and double-precision convergence control (10^{-7} tolerance limit).

3. Results and Discussion

The method section derives different forms of the LR equations for forensic evidence assessment when the data category is a continuous data distribution type, either following the normal distribution or t-distribution, which are the two most useful distributions used in scientific evidence assessment at a crime scene. As a numerical example, consider a questioned glass fragment on the defendant's coat has a scaled refractive index of $x = 110$ and the maximum scaled uncertainty in the measurement, " Δ " ranges from 20 to 0.1 with suitable increments. A set of control glass fragments at the crime scene is measured similarly, giving a mean scaled refractive index of $\bar{X} = 100$ with a scaled standard deviation of $S = 10$. Let us also assume that the data regarding the refractive indices of glass fragments sampled from a target population representing the environmental background naturally follow a normal distribution with $\mu = 70$ and $\sigma = 20$. The benchmark numerical values ($x = 110, \bar{X} = 100, S = 10, \mu = 70, \sigma = 20$) correspond to the normalised standard deviation distance profiles (Z-scores) used in practice during forensic casework. In addition to this, an additional parameter sensitivity range with different mean offset and standard deviation ratio has been analysed to verify the parameter independence.

Based on the competing propositions H_1 that states the questioned glass fragment on the suspect is from the crime scene and H_2 which states the questioned glass fragment on the suspect is from the environment. The likelihood ratio for this evidence, based on these measurements of the glass fragment, is shown in Figure 1 as

a function of error percentage of $100\Delta/x$. A normal distribution is assumed for both hypotheses, representing the numerator and denominator of the LR when performing the calculations for Figure 1. The exact method (exact probability integration approach) using Eq. 2 and the approximate method (probability density function approximation) using Eq. 5 are compared. This figure is based on finding the magnitude of LR for decreasing values of " Δ " using Eq. 2 and is compared with one without a known " Δ " as in Eq. 5 or 8.

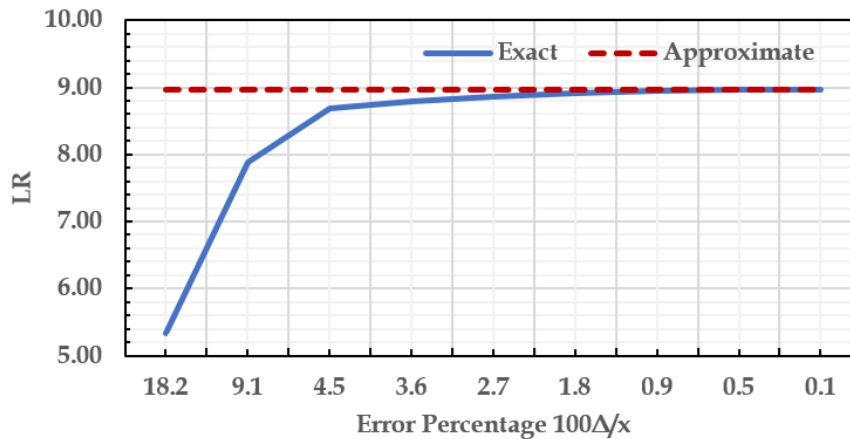


Figure 1. Likelihood ratio as a function of measurement error based on normal distribution

An examination of this figure shows that the lower the measurement uncertainty under study, " Δ ", LR calculated by the exact method using Eq. 2 converges towards a value of 8.96. When the maximum error of measurement approaches 2% of the actual measurement, the approximate method can be used without the need to know the actual value of the uncertainty. Both methods coincide when the error of measurement is equal to or less than 2% of the actual measurement. In fact, from the perspective of probability calculation accuracy, the variance in the probability ratio decreases as the error value " Δ " decreases. Despite the noticeable change in the numerator and denominator values of the LR , the ratio between them remains of the same order of magnitude and even converges to a constant value.

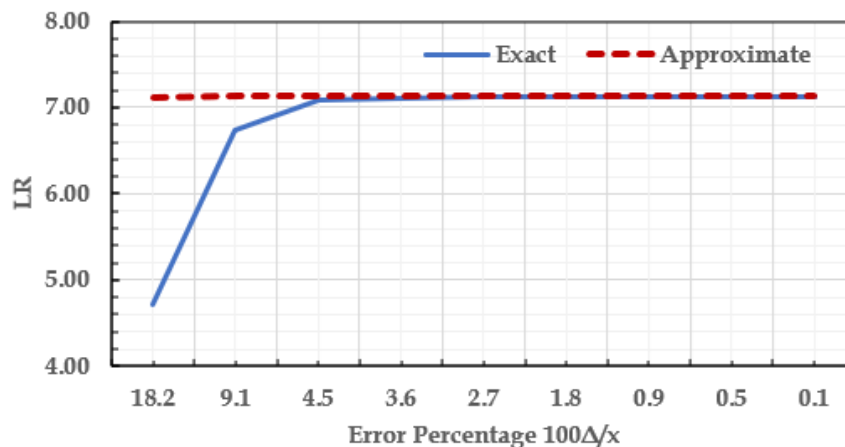


Figure 2. Likelihood ratio as a function of measurement error based on t-distribution ($DOF = 2$)

In the same manner, Figure 2 shows the variation of the LR with the measurement error based on a t-distribution with a degree of freedom $DOF = 2$. A degree of freedom equal to two means the sample size of the control

sample consists of three glass fragments. A sample consisting of three specimens is considered one of the most common and realistic types of samples found at crime scenes. The t-distribution is assumed for nominator and the normal distribution is used for the denominator of the likelihood ratio for the calculation of Figure 2. A community emerging from the environment is typically large and behaves like a normal distribution, particularly when analysing scientific evidence.

The exact method using Eq. 2 and the approximate method using Eq. 5 based on the t-distribution are compared in Figure 2. Again, by inspection of this figure, as the uncertainty in the measurement “ Δ ” under question decreases, LR calculated by the exact method converges towards a value of 7.13. When the maximum error of measurement approaches 4.5% of the actual measurement, the approximate method can be used without the need to know the actual value of the uncertainty in the measurement. Both methods coincide when the error of measurement is equal to or less than 4.5% of the actual measurement in the case of the t-distribution with two degrees of freedom. Here again, the variation across this figure is insignificant for a small value of error “ Δ ” and indeed converges to a constant value.

When the degree of freedom based on the t-distribution increases from 2 to 4, 8, 16 and 32, the behaviour of the likelihood ratio as a function of maximum error (Eq. 2) approaches the normal behaviour as illustrated in Figure 3. Examination of this figure shows that the normal distribution can be used instead of the t-distribution with good precision when the degree of freedom is more than 16. In contrast, when the degree of freedom is less than 8, significant differences in the value of the LR are observed.

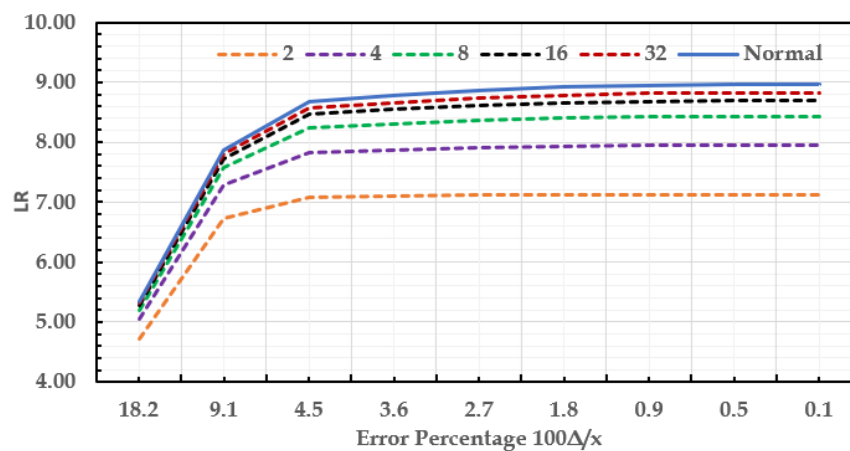


Figure 3. Likelihood ratio as a function of measurement error with different degrees of freedom

The approximate method using Eq. 5 for the t-distribution is applied for different degrees of freedom. Figure 4 shows the value of the LR for each degree of freedom. Inspection of this figure shows that the approximate value of the LR increases as the degrees of freedom increase. Furthermore, when the degrees of freedom become large, the value of the likelihood ratio approaches the equivalent value obtained using a normal distribution.

Finally, Figure 5 shows the convergence error “ Δ ” as a percentage of the true measurement “ x ” when the two methods give the same value of the likelihood. As shown in Figure 1 and Figure 5, both methods coincide when the error of measurement is equal to or less than 2% of the actual measurement if the distribution follows the normal shape. As shown in Figure 2 and Figure 5, both methods coincide when the error of measurement is equal to or less than 4.5% of the actual measurement if the distribution follows the t-shape with two degrees of freedom. The convergence error decreases as the degrees of freedom increase, as shown in Figure 5, reflecting the additional variability associated with a small sample size.

3.1. Validation Based on Real Casework Dataset in Forensic Science

To validate the theoretical limitations and test the applicability of the proposed methodology in practical forensic conditions. The analysis of a real casework dataset obtained by means of forensic refractive index

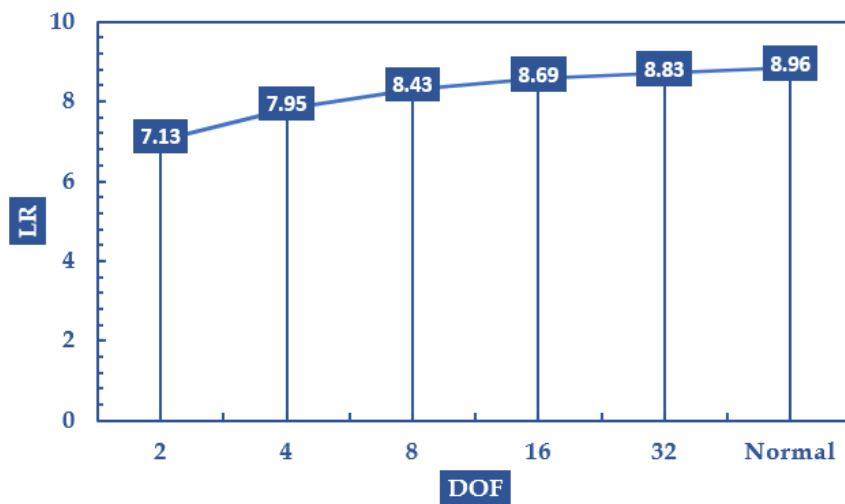


Figure 4. The LR as a function of degrees of freedom using the approximate method

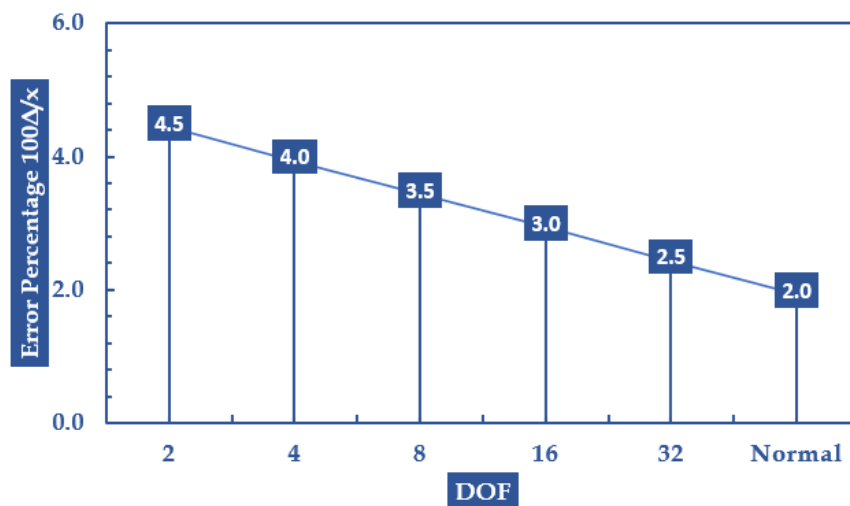


Figure 5. Variation of convergence percentage error with different degrees of freedom

(RI) measurements of glass samples using the $GRIM$ system was carried out ([17]; [5]). In the particular case considered here, a glass fragment collected from a broken window as a control item had a mean RI value of $\bar{X} = 1.51862$ with a standard deviation $S = 0.00012$ for $N = 3$ pieces (" DOF " = 2). One questioned glass fragment obtained from the suspect's clothing had RI of $x = 1.51878$. The experimental measurement uncertainty of $GRIM3$ calibrated procedure equals $\Delta = \pm 0.00005$, with a relative measurement error ($100\Delta/x$) of approximately 0.0033%. The parameters of the background reference population for the floating window glass were $\mu = 1.51810$ and $\sigma = 0.00240$. The calculation of the likelihood ratio by exact probability integration (Eq. 2), taking into account the experimental uncertainty interval ($\Delta = \pm 0.00005$), gave " LR " = 7.14. Using the probability density function approximation (Eq. 5) without explicitly considering the measurement uncertainty Δ gave " LR " = 7.15. Given the fact that the uncertainty of the operation measurement for modern forensic instrumentation (0.0033%) is significantly lower than the set thresholds for convergence (2% in case of normal distribution and 4.5% for Student's t -distribution when " DOF " = 2). Both approximate and precise calculations

resulted in almost the same values of evidential strength. In this way, the practical application proved that for highly accurate analysis equipment used in regular forensic labs, approximation is equally reliable.

3.2. Clarification and Practical Framing of Measurement Uncertainty

To avoid confusion with respect to the use of measurement uncertainty within the scope of forensic applications, it is important to distinguish between two distinct parts:

Instrumental Measurement Uncertainty (u_{inst} or Δ): This is the fundamental, fixed precision of the lab equipment and method of analysis (e.g., the standard error of calibration for the GRIM3 system according to ASTM E1967 of $\Delta = \pm 0.00005$). In the routine work of a forensic laboratory, it is a fixed value defined during the procedure validation and cannot be arbitrarily modified by the analyst.

Sampling Variability ($S/\sqrt{N}$): It is the standard error of the mean of the sample caused by the heterogeneity of the collected pieces of glass from the crime scene.

In this study, the examination of the quantity $100\Delta/x$ on different values in the figures and in simulations was not meant to allow the analyst to manually modify the precision of the instrument. On the contrary, it was a way to test the sensitivity of the method and set the benchmarks for different analytical instruments (e.g., GRIM3, ICP – MS, and SEM – EDS) and their different historical calibration standards. Indeed, the practical usefulness of the suggested 2% (normal distribution) and 4.5% (Student's t-distribution with $DOF = 2$) limits lies in the possibility of providing the analyst with a criterion for when he or she should make the decision whether to employ an approximate probability density function (PDF) method (Eq. 5) and do not have any risks of numerical bias and lack of evidential integrity, or whether an exact integral approach (Eq. 2) is required due to high instrumental uncertainty of the laboratory instrument.

3.3. Generalizability for Other Continuous Forensic Modalities

Though the refractive index (RI) of glass fragments is used as a standard continuous forensic sample, the obtained convergence limits (2% for Normal and 4.5% for Student's t-distribution, " $DOF = 2$ ") may be directly generalised for other forensic data of a continuous nature that is dependent on physical instrumentation. Thus, for example, elemental analysis by means of ICP – MS or SEM – EDS (e.g., concentrations of elements in forensic soil, paint, or gunshot residue samples), wherein standard relative analytical uncertainty values are usually within the range of 0.1% to 1.5%, the approximate PDF approach will still be entirely applicable.

3.4. "The Rare Type" Issue and its Legal Consequences

In forensic casework practice, one of the key issues that appear concerns the situation when the physical evidence (e.g., a fragment of glass) is a "rare type" and is an extreme outlier from the reference population database. Likelihood ratio (LR) of continuous data in probability density modelling depends critically on the tail behaviour of the presumed background distribution.

Behaviour of Normal vs. Student's t-distribution in Tail: When assuming normality, the probability density function $f(x)$ decreases exponentially fast toward zero in the tails. In other words, in the case of rare fragments ($x \gg \mu$), the denominator of the LR formula converges quickly to zero. In such an extreme situation, using the approximate PDF approach may cause overestimation of LR (weight of evidence) because of mathematical inaccuracies in the approximation of a very small denominator. On the contrary, the Student's t-distribution has heavy tails and decays polynomially. The heavy-tailedness of the background helps avoid zero-frequency issues and artificial increases in LR.

Influence of the Approximate Approach on the Weight of Evidence: For common or moderately rare samples, the approximate and exact approaches provide practically equal LR values (as illustrated in the range of 2% and 4.5% thresholds). However, in the case of the extremes, the approximate PDF approach tends to slightly increase the LR compared to the exact integration approach when normality is assumed, while it maintains stability when a Student's t-distribution (with low degrees of freedom, e.g., $DOF = 2$) is implemented.

Legal Recommendation and Admissibility of Evidence: In legal terms, the overestimation of evidential value against a suspect goes against the rule of conservation. Therefore, to secure legal admissibility in court

practice related to rare glass types. Heavy-tailed distributions (Student's t-distribution) should be used to model the reference population in the presence of outliers. In the case of extreme outliers and near-zero background density, one should use the exact integration approach rather than an approximate one.

4. Conclusions

This study has explored in a systematic manner the limits of operation where exact numerical integration can be distinguished from point *PDF* approximations for the calculation of *LR* in the presence of measurement uncertainty. In this study, which combines the analytical error calculations with numerical simulation and real forensic glass *RI* data that conforms to *ASTME1967*(*GRIM3*), the following conclusions are arrived at:

Thresholds of Convergence and Evidence Quality:

For the *PDF* approximation in terms of points, the very high degree of accuracy as an alternative to integrating probability is ensured if the instrumental measurement error (Δ) is low in comparison with the standard deviation of the background distribution. For instance, relative measurement errors not greater than 2% in a Gaussian distribution and 4.5% in a t-distribution ("*DOF*" = 2) result in insignificant discrepancy (< 0.01%).

Instrumental Standard Uncertainty v. Sampling Variation:

The most important point to consider is that the instrumental standard uncertainty (u_{inst} or Δ), which is the basis of defining the integration range, should not be confused with sampling variation ($S/\sqrt{N}$), which affects the estimation of parameters. Current analytical laboratory instruments, such as *GRIM3* systems, which have a typical relative uncertainty of 0.0033%, and other *ICP – MS* and *SEM – EDS* techniques, function safely within the established convergence range, making numerical integration unnecessary.

Implementation and Compliance with ISO/IEC 17025:

The limit values that have been derived in the current work are considered empirical guidelines in an operational sense rather than universal physical principles. In the context of accredited laboratories that function based on the guidelines of *ISO/IEC17025*, the limits provide a universal screening tool that would determine whether or not closed-form *PDF* evaluation is safe or not.

Boundary Anomalies and the "Rare Type":

In the case of extreme continuous evidence where the observed trace lies at the outer tails of the background distributions (a "rare type" situation), the conventional continuous density model may yield excessively high or even infinitely large *LR* values. In this situation, it is necessary for those using such models to couple their use of *PDF* estimates with empirical probability bounds based on physical tolerance levels.

REFERENCES

1. Adam, C. (2011). Essential mathematics and statistics for forensic science. John Wiley & Sons.
2. Aitken, C. G. G., & Stoney, D. A. (1991). The Use Of Statistics In Forensic Science. CRC Press. <https://doi.org/10.1201/b12618>
3. Aitken, C. G. G., Taroni, F., & Bozza, S. (2004). Statistics and the evaluation of evidence for forensic scientists (Vol. 26). Wiley, Chichester.
4. Aitken, C., Taroni, F., & Bozza, S. (2024). The Role of the Bayes Factor in the Evaluation of Evidence. Annual Review of Statistics and Its Application, 11(1), 203–226. <https://doi.org/10.1146/annurev-statistics-040522-101020>
5. ASTM E1967-19 (2019). (2019). Test Method for the Automated Determination of Refractive Index of Glass Samples Using the Oil Immersion Method and a Phase Contrast Microscope. ASTM International. <https://doi.org/10.1520/E1967-19>
6. Buckleton, J. S., Bright, J.-A., & Taylor, D. (2018). Forensic DNA evidence interpretation. CRC Press.
7. Evett, I., & Weir, B. S. (1998). Interpreting DNA evidence: statistical genetics for forensic scientists. (No Title).
8. Lagzini, A., El Maroufy, H., Merbouha, A., & El Omari, M. (2025). Bayesian and Likelihood Inference for the SIR Model Using Skellam's Distribution with Real Application to COVID-19. Statistics, Optimisation & Information Computing, 15(4), 2397–2416. <https://doi.org/10.19139/soic-2310-5070-2993>
9. Leegwater, A. J., Meuwly, D., Sjerps, M., Vergeer, P., & Alberink, I. (2017). Performance study of a score-based likelihood ratio system for forensic fingerprint comparison. Journal of Forensic Sciences, 62(3), 626–640.
10. Lucy, D. (2013). Introduction to statistics for forensic scientists. John Wiley & Sons.
11. Mathews, P. G. (2004). Design of Experiments with MINITAB. Quality Press.
12. Mohamed Amine, B., & Mohammed, K. (2025). Analysis of uncertainty in the Leontief model by interval arithmetic. Statistics, Optimisation & Information Computing, 13(5), 2011–2026. <https://doi.org/10.19139/soic-2310-5070-2279>
13. Morrison, G. S. (2011). Measuring the validity and reliability of forensic likelihood-ratio systems. Science & Justice, 51(3), 91–98.

14. Neumann, C., Evett, I. W., & Skerrett, J. (2012). Quantifying the weight of evidence from a forensic fingerprint comparison: a new paradigm. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 175(2), 371–415.
15. NIST. (2020). National Institute of Standards and Technology (NIST), 2020. URL: <https://www.nist.gov/cyberframework>.
16. Ramos, D., Gonzalez-Rodriguez, J., Zadora, G., & Aitken, C. (2013). Information-theoretical assessment of the performance of likelihood ratio computation methods. *Journal of Forensic Sciences*, 58(6), 1503–1518.
17. Ramos, G. Z. A. M. D., & Aitken, C. (2014). *Statistical Analysis in Forensic Science: Evidential Value of Multivariate Physicochemical Data*.
18. Risan, H. K., & Al-Azzawi, A. A. (2025). Forensic analysis for the effect of Fourier series terms on the response of a structure to a periodic force. *International Journal of Forensic Engineering*, 5(3), 228–244.
19. Risan, H. K., Al-Azzawi, A. A., & Al-Zwainy, F. M. S. (2024). *Statistical Analysis for Civil Engineers*. Elsevier.
20. Robertson, B., Vignaux, G. A., & Berger, C. E. H. (2016). *Interpreting evidence: evaluating forensic science in the courtroom*. John Wiley & Sons.
21. Seliem, M. M., Kamel, A. R., Taha, I. M., & Abu El-Nasr, M. M. (2024). A Note on Prior Selection in Bayesian Estimation. *Statistics, Optimisation & Information Computing*, 13(2), 795–806. <https://doi.org/10.19139/soic-2310-5070-1752>
22. Taroni, F., Biedermann, A., Bozza, S., Garbolino, P., & Aitken, C. (2014). *Bayesian networks for probabilistic inference and decision analysis in forensic science*. John Wiley & Sons.
23. Willis, S. M., McKenna, L., McDermott, S., O'Donnell, G., Barrett, A., Rasmusson, B., Nordgaard, A., Berger, C. E. H., Sjerps, M. J., & Lucena-Molina, J. (2015). ENFSI guideline for evaluative reporting in forensic science. European Network of Forensic Science Institutes.