Closed-Form Solution of the Unit Normal Loss Integral in 2 Dimensions, with Application in Value-of-Information Analysis

Tae Yoon Lee®, Paul Gustafson®, and Mohsen Sadatsafavi®, for the Canadian Respiratory Research Network

The unit normal loss integral (UNLI) has wide applicability in decision analysis and risk modeling, including as a solution for computation of various value-of-information (VoI) metrics. However, one limitation of the UNLI has been that its closed-form solution is available for only 1 dimension and thus can only be used for comparisons involving 2 strategies (where it is applied to the scalar incremental net benefit). We derived a closed-form solution for the 2-dimensional UNLI by the integration by parts technique. This enables the extension of the UNLI method to 3-comparison problems.

We implemented this approach in R as part of the predtools package (https://github.com/resplab/predtools/) and verified the accuracy of this implementation via Monte Carlo simulations. A case study based on a 3-arm clinical trial was used as an example for VoI analysis. Methods based on the closed-form solutions for the UNLI can now be extended to 3-decision comparisons, taking a fraction of a second to compute and not being subject to Monte Carlo error.

Highlights

- The unit normal loss integral (UNLI) is widely used in decision analysis and risk modeling, including in the computation of various value-of-information metrics, but its closed-form solution is only applicable to comparisons of 2 strategies.
- We derive a closed-form solution for 2-dimensional UNLI, extending the applicability of the UNLI to 3-strategy comparisons.
- Such closed-form computation takes only a fraction of a second and is free from simulation errors that affect the hitherto available methods.
- In addition to the relevance in 3-strategy model-based and data-driven decision analyses, a particular application is in risk prediction modeling, where the net benefit of a classifier should always be compared with 2 default strategies of treating none and treating all.

Keywords

decision analysis, risk prediction modeling, cost-effectiveness, uncertainty, value of information

Date received: August 16, 2022; accepted: March 29, 2023
The unit normal loss integral (UNLI), first discussed by Raiffa and Schlaifer in the 1960s, has emerged in various ways in decision analysis and risk modeling, including in value-of-information (VoI) analysis. 1 VoI analysis is a set of concepts and methods rooted in decision theory that quantifies the expected utility loss due to uncertainty associated with decisions. 2 VoI has been applied in decision analysis across areas including health technology assessment, 3 environmental risk analysis, 4 and clinical prediction modeling. 5 VoI calculations are often carried out using Monte Carlo (MC) simulations, either through repeated sampling of uncertain input parameters in model-based evaluations or via bootstrapping in data-driven analyses. 6

An overview of UNLI-based methods for VoI computations is provided by Wilson. 7 The UNLI is closely related to the mean of the truncated normal distribution. The exact definition of the UNLI has varied slightly in different publications. 7–9 Here we use the following definition. Suppose Y has a normal distribution with mean μ and variance σ². With \( \phi(y) = \frac{e^{-\frac{y^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \) and \( \Phi(y) = \int_{-\infty}^{y} \phi(t)dt \) denoting, respectively, the probability density and cumulative distribution functions of the standard normal distribution, the UNLI can be defined as

\[
\text{UNLI} = E(\max(Y, 0)) = \int_{-\infty}^{\infty} \frac{e^{-\frac{(y-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dy
\]

\[
= \int_{-\infty}^{\infty} \frac{y}{\sigma} \phi \left( \frac{y-\mu}{\sigma} \right) dy = \mu \left[ 1 - \Phi \left( \frac{\mu}{\sigma} \right) \right] + \sigma \Phi \left( \frac{\mu}{\sigma} \right).
\]

A typical application in VoI analysis is the computation of the expected value of perfect information (EVPI). 1,10 EVPI is the expected gain in net benefit (NB) when uncertainty in the evidence underlying the decision is completely resolved. 7 When comparing 2 strategies (e.g., use of a new medication versus continuing with standard of care), the outcome of a probabilistic decision analysis can be summarized as a distribution of the incremental NB between the 2 strategies. If this quantity has a normal distribution, then the EVPI can be expressed as a closed-form solution using the 1-dimensional UNLI, as outlined above. 1,10

This approach for EVPI calculation is applicable to both model-based and data-driven evaluations. The UNLI method has been extended to other VoI metrics, such as expected value of partial perfect information 11 and expected value of sample information. 12 Such solutions are computationally feasible and free from MC error, albeit requiring the assumption of normality to hold. However, in many practical decision analyses, there are more than 2 strategies that are compared, making the closed-form UNLI not readily applicable. Approximate methods have been suggested for more than 2 strategies. For example, Jalal et al. 8 used the UNLI as an approximate solution to multiple comparisons by segmenting the joint probability space of input parameters into adjacent pieces to turn the problem into a sum of 1-dimensional evaluations.

To the best of our knowledge, no closed-form expression for the UNLI has been proposed for higher dimensions. In this work, we derive a closed-form solution for the UNLI for 2 dimensions, enabling the extension of this method to comparisons of 3 strategies. We perform a simulation study to verify the numerical accuracy of our implementation and show its utility in a case study involving EVPI calculation for a data-driven decision analysis based on a 3-arm clinical trial.

Closed-Form Solution

Suppose we have 3 strategies of interest, with one of them labeled as the reference strategy (the choice of which strategy being designated as the reference has no bearing on the computation). The incremental NBs of the two alternative strategies compared with the reference strategy are denoted by \( Y_1 \) and \( Y_2 \). Further, suppose that our knowledge of \( Y = (Y_1, Y_2) \) can be expressed as a bivariate normal distribution with mean \( (\mu_1, \mu_2) \), variance \( (\sigma_1^2, \sigma_2^2) \), and correlation coefficient \( \rho \). The expected NB under perfect information is the expectation of \( \max(Y, 0) \).

Following the derivations provided in the Supplementary Material Section 1, we arrive at the closed-form equation:
\[ E(\max(Y, 0)) = r_{1,2} + q_{1,2} + r_{2,1} + q_{2,1}, \]

with each term on the right-hand side defined below. In what follows, \( I\{\text{condition}\} = 1 \) if the condition is true and 0 otherwise; \( \Phi_2(x_1, x_2, \rho) \) is the cumulative density function of the standard bivariate normal with the upper limits and \( x_2 \) and correlation coefficient \( \rho \), and

\[
\delta_{i,j} = \begin{cases} 
1 & \text{if } \sigma_i - \rho \sigma_j > 0 \\
-1 & \text{if } \sigma_i - \rho \sigma_j < 0
\end{cases}
\]

\[
r_{i,j} = \mu_i \\
\left[ \Phi_2\left( \frac{\delta_{i,j}(\mu_j - \mu_i)}{\sqrt{\sigma_i^2 + \sigma_j^2 - 2\rho \sigma_i \sigma_j}} \right) - \Phi_2\left( \frac{-\delta_{i,j}(\mu_j - \mu_i)}{\sqrt{\sigma_i^2 + \sigma_j^2 - 2\rho \sigma_i \sigma_j}} \right) \right] \\
\text{and}
\]

\[
q_{i,j} = I\{(\sigma_i - \rho \sigma_j) \neq 0\} \mu_i \delta_{i,j} \Phi_2\left( \frac{\delta_{i,j}(\mu_j - \mu_i) - \sigma_i \sigma_j \sqrt{(1 - \rho^2)}}{\sigma_i^2 + \sigma_j^2 - 2\rho \sigma_i \sigma_j} \right)
\]

\[
\left( 1 - \Phi_2\left( \frac{-\delta_{i,j} \sigma_i^2 + \sigma_j^2 - 2\rho \sigma_i \sigma_j}{\sigma_i - \rho \sigma_j} \right) \right) \\
\delta_{i,j} \Phi_2\left( \frac{\delta_{i,j} \sigma_i^2 + \sigma_j^2 - 2\rho \sigma_i \sigma_j}{\sigma_i - \rho \sigma_j} \right)
\]

**Case Study**

The Optimal Therapy of Chronic Obstructive Pulmonary Disease (COPD) was a parallel-arm clinical trial of 3 inhaler therapies for patients with COPD. In this 3-arm study, 449 patients with COPD were randomized to receive single-inhaler \((n = 145)\), double-inhaler \((n = 156)\), or triple-inhaler \((n = 148)\) therapies. The trial duration was 12 mo. The study collected data, including monthly cost diaries, and functional scores measured by St. George’s Respiratory Questionnaire at baseline and 4 follow-up visits (at 4, 20, 36, and 52 wk).

The NB calculations in this case study closely follow the methods used in a previously published data-driven economic evaluation of this trial. The functional scores were converted to EQ5D utilities using validated algorithms. With the quality-adjusted life years (QALYs) as the health outcome of interest, the single-inhaler strategy dominated the double-inhaler strategy, and the incremental cost-effectiveness ratio of the triple-inhaler therapy versus the single-inhaler therapy was $4,042–2,678)/[(0.7217–0.7092)QALY] = $243,180/QALY gained (in 2006 Canadian dollars). As such, the single-inhaler therapy was the optimal strategy at the willingness-to-pay value of $50,000/QALY. However, a bootstrap-based probabilistic sensitivity analysis demonstrated uncertainty in the results: the single-inhaler therapy was the optimal strategy in 80% of the bootstraps.

We calculated the EVPI for this evaluation based on such individual-level data from the trial, comparing the bootstrap-based approach with the proposed UNLI method. At a given willingness-to-pay value, the NB was

**Implementation and a Simulation Study**

We implemented this method in R (mu_max_trunc_bvn function in the predtools package - https://github.com/reslab/predtools). To verify the accuracy of our implementation, we conducted a simulation study by comparing it with large-scale MC integrations \((N = 100,000)\). We examined 252 permutations of the parameters of the bivariate distributions characterized by a factorial design for the following variables: \( \mu_1, \mu_2 = \{-2, 0, 2\}, \sigma_1^2, \sigma_2^2 = \{1, 3\} \), and \( \rho = \{-0.75, -0.50, \ldots, 0.50, 0.75\} \). Results show that the difference between the closed-form and MC solutions falls within the range of the MC standard error (Supplementary Material Section 2).
calculated as QALY multiplied by willingness-to-pay minus total costs for each patient. Then, for each arm, we calculated the average NB. Seven percent of costs and utility values were missing. Similar to the approach used in the original evaluation, we imputed the missing values using multiple imputation by chained equation with predictive mean matching. For the bootstrap-based approach, imputation was embedded within bootstrapping, such that each iteration of the MC simulation involved one imputation and generation of a single bootstrapped sample. This was repeated 1,000 times. For the UNLI method, we generated 10 imputed data sets and pooled the mean and covariance matrix estimates of incremental NBs using the Rubin’s rule. Next, taking the single-inhaler therapy as the reference, the incremental NBs of double- and triple-inhaler therapies were parameterized as a bivariate normal distribution. For example, at a willingness-to-pay of $50,000/QALY, the pooled parameter values were $\mu_1 = -4.734$, $\mu_2 = -2.668$, $\sigma_1 = 4.678$, $\sigma_2 = 4.645$, and $\rho = 0.50$, where subscripts 1 and 2 indicate the incremental NB of double- and triple-inhaler therapies, respectively, compared with single-inhaler therapy.

EVPI values were similar between the closed-form UNLI and bootstrap methods with the mean relative absolute error of 3% (Figure 1). At a willingness-to-pay of $50,000/QALY, the EVPI values were $1,019 for the closed-form UNLI method and $976 (MC standard error: 61) for the bootstrap method. At a willingness-to-pay of $100,000/QALY, the corresponding values were $1,570 and $1,543 (MC standard error: 107).

**Discussion**

We have proposed a closed-form solution for the 2-dimensional version of the UNLI, enabling the extension of existing methods based on this integral to 3-strategy decision problems, with particular application in VoI analysis. Provided the assumption of the bivariate normality of the incremental NBs is ascertained, the UNLI method provides a solution that takes a fraction of a second to compute and is not subject to the MC error inherent in simulation-based methods. The R code for this method is provided in the predtools package, along with a tutorial (https://reslab.github.io/predtools/articles/UNLI2D.html).

In a case study, we compared the UNLI method with the conventional bootstrap-based approach for EVPI computations for a data-driven economic evaluation based on data from a 3-arm clinical trial. We found that the results from the bootstrap-based approach and the UNLI-based approach were very close. As the UNLI-based approach is computationally faster and free from MC error, in similar contexts, presenting the UNLI-based results in VoI analysis would help with the reproducibility of results.

A key condition for the validity of UNLI-based computations is the normality assumption (bivariate normality in the 2-dimensional case). If this condition is violated, the UNLI-based approach could produce misleading results (in which case, more flexible methods should be considered). In data-driven evaluations, provided that the sample size is sufficient, the central limit theorem provides a general justification for the normality assumption with sufficient sample sizes. For model-based evaluation, the appropriateness of the normality assumption must be scrutinized, as has been demonstrated in the previous applications of the 1-dimensional UNLI method.

In addition to decision-analytic problems, the 2-dimensional UNLI can be particularly relevant for the recently proposed VoI analysis in risk prediction modeling. The NB calculation for a risk prediction model is naturally a 3-decision problem, as a risk prediction model must be always compared with at least the 2 default strategies of “treating none” and “treating all.” Here again, the central limit theorem provides justification for the assumption of (bivariate) normality, for example, for the distribution of incremental NB of “treating all” and “model-based treatment” with “treating none” as the reference strategy. The Wald-type method proposed by Marsh et al. can be used to derive the moments of the bivariate normal distribution, enabling UNLI-based VoI analysis.
In this work, we focused on EVPI calculation to demonstrate the applicability of the 2-dimensional UNLI method. The 1-dimensional UNLI method has been expanded to other VoI metrics, including the expected value of partial perfect information and the expected value of sample information.7–8 Those algorithms can be feasibly modified to accommodate decision analyses with 3 strategies using the 2-dimensional UNLI method.

A natural question is whether the UNLI method can be extended to higher dimensions. We believe that there is no closed-form solution for the general m-dimensional case. The m-dimensional UNLI is closely related to the expression for the maximum of m-variate normal random variables. Arellano-Valle and Genton showed (corollary 4) that for m > 2, the probability density function of the maximum of m-variate normal random variables is no longer composed of univariate normal distribution functions.22 Under the assumption of independence, DasGupta provided a closed-form solution for the expectation of 3 normal random variables and asserted that such a solution is unlikely to exist for m > 3.23 Thus, it is not surprising that for a general setting, only bounds on the expectation, or infinite series expansions, rather than a closed-form solution, have been computed in the literature.24

Acknowledgments
We would like to thank the abstract reviewers for their suggestions and comments at the Society of Medical Decision Making 44th Annual North American Meeting and thank the editor and reviewers for their feedback, which improved the quality of this article.

ORCID iDs
Tae Yoon Lee https://orcid.org/0000-0001-6672-4317
Paul Gustafson https://orcid.org/0000-0002-2375-5006
Mohsen Sadatsafavi https://orcid.org/0000-0002-0419-7862

Supplemental Material
Supplementary material for this article is available online at https://doi.org/10.1177/0272989X231171166.

References
1. Schlaifer RO. Probability and Statistics for Business Decisions: An Introduction to Managerial Economics Under Uncertainty. New York: McGraw-Hill; 1959.
2. Howard RA. Information value theory. IEEE Trans Syst Sci Cybern. 1966;2(1):22–6.
3. Fenwick E, Steuten L, Knies S, et al. Value of information analysis for research decisions—an introduction: report 1 of the ISPOR value of information analysis emerging good practices task force. Value Health. 2020;23(2):139–50.
4. Borgonovo E, Cillo A. Deciding with thresholds: importance measures and value of information. Risk Anal. 2017;37(10):1828–48.
5. Sadatsafavi M, Yoon Lee T, Gustafson P. Uncertainty and the value of information in risk prediction modeling. Med Decis Making. 2022;42(5):661–71.
6. Sadatsafavi M, Marra C, Bryan S. Two-level resampling as a novel method for the calculation of the expected value of sample information in economic trials. Health Econ. 2013;22(7):877–82.
7. Wilson ECF. A practical guide to value of information analysis. Pharmacoeconomics. 2015;33(2):105–21.
8. Jalal H, Goldhaber-Fiebert JD, Kuntz KM. Computing expected value of partial sample information from probabilistic sensitivity analysis using linear regression metamodeling. Med Decis Making. 2015;35(5):584–95.
9. Hellermann R. Capacity Options for Revenue Management: Theory and Applications in the Air Cargo Industry. New York: Springer; 2006.
10. Schlaifer R. Introduction to Statistics for Business Decisions. New York: MacGraw-Hill; 1961.
11. Coyle D, Oakley J. Estimating the expected value of partial perfect information: a review of methods. Eur J Health Econ. 2008;9(3):251–9.
12. Willan AR, Pinto EM. The value of information and optimal clinical trial design. Stat Med. 2005;24(12):1791–806.
13. Aaron SD, Vandemheen K, Fergusson D, et al. The Canadian optimal therapy of COPD trial: design, organization and patient recruitment. Can Respir J. 2004;11(8):581–5.
14. Najafzadeh M, Marra CA, Sadatsafavi M, et al. Cost effectiveness of therapy with combinations of long-acting bronchodilators and inhaled steroids for treatment of COPD. Thorax. 2008;63(11):962–7.
15. Meguro JP. Elicitation of utility weights for a copd specific preference measure. Am J Respir Crit Care Med. 2006;173(3):A488.
16. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw. 2011;45:1–67.
17. Heymans MW, Eekhout I. Applied Missing Data Analysis with SPSS and (R)Studio. 1st ed. Amsterdam; 2019. Available from: https://bookdown.org/mwheymans/bookm/ [Accessed 5 July, 2022].
18. Briggs A, Claxton K, Sculpher M. Decision Modelling for Health Economic Evaluation. Illustrated ed. Oxford (UK): Oxford University Press; 2006.
19. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Mak. 2006;26(6):565–74.
20. Marsh TL, Janes H, Pepe MS. Statistical inference for net benefit measures in biomarker validation studies. *Biometrics*. 2020;76(3):843–52.
21. Sadatsafavi M, Lee TY, Wynants L, Vickers A, Gustafson P. Value of information analysis for external validation of risk prediction models. 2022. arXiv: 220803343.
22. Arellano-Valle RB, Genton MG. On the exact distribution of the maximum of absolutely continuous dependent random variables. *Stat Probab Lett*. 2008;78(1):27–35.
23. DasGupta A. A formula for the expected value of the maximum of three independent normals and a sparse high dimensional case. *Purdue University*; 2021. Available from: https://www.stat.purdue.edu/~dasgupta/orderstat.pdf [Accessed 3 December, 2022].
24. Ross AM. Computing bounds on the expected maximum of correlated normal variables. *Methodol Comput Appl Probab*. 2010;12(1):111–38.