AIMS Mathematics, 11(6): 17320–17353.
DOI: 10.3934/math.2026709
Received: 01 April 2026
Revised: 24 May 2026
Accepted: 01 June 2026
Published: 15 June 2026

https://www.aimspress.com/journal/Math

Research article

Prior-free probabilistic interval estimation for constrained parameters

Hezhi Lu[1,2,*] **and Qijun Wu**[1]

[1] School of Economics and Statistics, Guangzhou University, Guangzhou 510006, China
[2] Lingnan Research Academy of Statistical Science, Guangzhou University, Guangzhou 510006, China

* **Correspondence:** Email: luhz@gzhu.edu.cn.

Abstract: In applied statistics, constructing exact confidence intervals for constrained normal means and constrained Poisson means are two long-standing challenges. Most existing inferential methods assume that the nuisance parameters in constrained models are known constants, which is often impractical. When nuisance parameters are unknown, Bayesian intervals fail to guarantee nominal coverage. To address these issues, this work developed a valid prior-free inferential model approach for normal and Poisson distributions with unknown nuisance parameters. For the constrained normal case, we constructed an inferential model interval that exactly attains the prespecified coverage probability. For the constrained Poisson case, we first proposed an inferential model interval that achieves exact frequentist coverage control; however, owing to the discreteness of the Poisson distribution, this interval is conservative. We then improved it by introducing random weighting, yielding a nonrandomized inferential model method. Simulation studies showed that the inferential model interval achieves exact coverage under the constrained normal model, but is conservative in constrained Poisson inference. In contrast, the nonrandomized inferential model interval attains coverage closest to the nominal level by shortening the interval length. Finally, two neutrino datasets from high-energy physics were used to demonstrate the advantages of the proposed new methods over Bayesian approaches.

Keywords: constrained parameter; confidence interval; improper prior; inferential model; predictive random set
Mathematics Subject Classification: 62F30, 62P35

1. Introduction

In statistical inference, parameters often have known constraints such as nonnegativity or boundedness. Effectively quantifying uncertainty and constructing interval estimates for such

parameters remains a persistent challenge for both frequentist and Bayesian methods. These problems are particularly common in scientific fields such as high-energy physics, astronomical observation, and environmental monitoring. For example, high-energy physicists have been particularly interested in two inference problems over the past two decades: inference about a nonnegative mean subject to Gaussian measurement error, and inference about a Poisson signal rate in the presence of background contamination. In this paper, we aim to develop more efficient inference methods for such constrained parameter problems, which are essentially based on an inferential model (IM) framework.

Numerous researchers have investigated constrained parameters under two distinct modeling assumptions. Under the first assumption, for the normal case, suppose that X is the measurement of a nonnegative quantity, θ, with a Gaussian error distribution. For simplicity, the variance is set to $\sigma^2 = 1$, which yields the probability model $X \sim N(\theta, 1)$ subject to the constraint $\theta \geq 0$. For the Poisson case, let $S \sim Poisson(\lambda)$ denote the number of signal events and $B \sim Poisson(\varepsilon)$ denote the number of background events, with a known background rate ε and an unknown signal rate λ. Assuming the independence of S and B, only their sum $X = B + S$ is observed, which follows a Poisson distribution with mean $\theta = \varepsilon + \lambda$. Inference on λ is therefore equivalent to inference on θ, subject to the physical constraint $\theta \geq \varepsilon$. When an inference is performed on a constrained parameter, traditional Neyman confidence intervals (CIs) can yield empty intervals or exhibit unsatisfactory coverage properties when the data are near or beyond the constraint boundaries. Feldman and Cousins [1] and Mandelkern and Schultz [2] proposed unified approach-based CIs for bounded parameters. Giunti [3] demonstrated that the upper bound of the unified CI decreases as ε increases, and introduced a new ordering principle to modify the unified approach. Modest refinements have been proposed by Mandelkern [4], Fraser et al. [5], and Zhu [6]. However, although these frequentist CIs can ensure a preset coverage probability, their frequentist coverage probabilities substantially exceed the nominal confidence level.

Bayesian inference provides a more credible interval estimation method than the unified interval does. However, this construction method yields unphysically short CIs near the boundary values of the constraint conditions. For example, Roe and Woodroofe [7] adopted a Bayesian approach based on a uniform improper prior to constructing a credible belt for constrained parameters. Zhang and Woodroofe [8] reported that Bayesian CIs based on a uniform prior still maintain a high coverage probability. In fact, Bayesian methods that use noninformative priors typically lack a clear frequentist interpretation and are sensitive to the choice of prior specifications. To eliminate the influence of prior distributions on inferential results, Leaf and Liu [9] and Lu et al. [10] proposed IM inferential methods for constrained parameters and developed a postdata predictive measure of uncertainty for unknown parameters. Moreover, Chen et al. [11] proposed a new fiducial interval for the constrained Poisson parameter. Although the IM and fiducial intervals can improve the conservative coverage problem of frequentist and Bayesian intervals, these methods are still not optimal for constrained parameter inference. Therefore, the optimal CIs of constrained parameters are still in progress.

Many research findings have been reported regarding the assumption of the first model; however, Zhang and Woodroofe [8] reported that the true values of the nuisance parameters σ^2 and ε in normal and Poisson models are difficult to determine in practical scenarios. They thereby proposed a more general theoretical model: In the normal case, suppose that independent random variables $X \sim N(\theta, \sigma^2)$ and $W \sim \sigma^2\chi^2_r$ are observed, where $\theta \geq 0$ and $\sigma > 0$ are unknown parameters and r is a positive integer. In the Poisson case, suppose that we observe $X = B + S$ and W, where B, S, and W are independent random variables for which $B \sim Poisson(\varepsilon)$, $S \sim Poisson(\lambda)$, and $W \sim Poisson(m\varepsilon)$, where $\lambda, \varepsilon \geq 0$

are unknown parameters and $m > 0$ is known. The goal is to develop interval estimation methods for the constrained parameters of interest, θ and λ, in the context of normal and Poisson distribution cases when the nuisance parameters σ^2 and ε are unknown.

The literature on the assumption of the second model is relatively scarce. On the basis of our experience, only Zhang and Woodroofe [8] systematically investigated the construction of credible sets for parameters under constrained parameter spaces via improper priors. Our simulation studies reveal that Bayesian credible intervals for the constrained parameter are unstable and fail to guarantee the nominal coverage probability. The IM framework is highly suited to scenarios with constrained parameter spaces. It directly integrates the parameter constraints with predictive random sets (PRSs), thereby addressing the challenges that such spaces pose to frequentist and Bayesian methods. More importantly, grounded in Dempster–Shafer belief function theory (Dempster–Shafer [12, 13]), the IM yields frequentist-calibrated inferences without prior information; therefore, this paper focuses on the second theoretical model and proposes more robust and interpretable IM-based CIs for constrained parameters.

The innovations and contributions of this paper are as follows. First, we extend statistical models that require the nuisance parameters to be known constants to models where the nuisance parameters are unknown but auxiliary observations are available. The new framework is more general and more practical. Second, existing IM approaches (e.g., Leaf and Liu [9]; Lu et al. [10]) assume known nuisance parameters. Consequently, they use only a single auxiliary variable to link the data and the parameter of interest. In contrast, after obtaining auxiliary observations, we introduce new auxiliary variables to link the auxiliary data and the nuisance parameters, thereby constructing a more precise IM joint association model. Third, when incorporating parameter constraints, we propose a minimal non-empty PRS for the parameter of interest. Compared with the PRS stretching approach (Leaf and Liu [9]), the new method is simpler and more computationally efficient. Fourth, we establish an exact finite-sample coverage theory for the proposed IM method. For the constrained Poisson mean, we further extend the random weighting method (Lu et al. [10]). The resulting nonrandomized IM method outperforms both the IM and Bayesian methods in terms of coverage probability and interval length. Notably, other interval construction methods, such as profile likelihood and bootstrap, are also available. However, the profile likelihood method tends to undercover near parameter boundaries and requires cumbersome handling of nuisance parameters. The bootstrap method is sensitive to the resampling scheme and often fails to achieve the nominal coverage level when the parameter is close to the boundary. Moreover, the IM plausibility function offers an intuitively interpretable measure of uncertainty for each parameter value. This is a unique feature not shared by Bayesian intervals, profile likelihood intervals, or bootstrap intervals.

The remainder of the paper is structured as follows: Section 2 reviews the models and Bayesian methods from Zhang and Woodroofe [8]. Section 3 introduces the basic framework of the IM and proposes our IM-based inference methods for constrained parameters. In Section 4, we compare the coverage performance of the new method with that of Bayesian approaches through simulation studies. In Section 5, we give two real data analysis for illustration of our methods. Discussion and Conclusions are presented in Sections 6 and 7, respectively.

2. Existing interval estimation methods

2.1. Inference for nonnegative quantities with Gaussian errors

Suppose that X is the measurement of a nonnegative quantity θ, following a Gaussian error distribution $N(\theta, \sigma^2)$ with mean θ and variance σ^2, where θ satisfies the constraint $\theta \geq 0$ and σ^2 is a nuisance parameter. In the study of Zhang and Woodroofe [8], the model further assumes that the observed random variables $X \sim N(\theta, \sigma^2)$ and $W \sim \sigma^2\chi_r^2$ are mutually independent, where $\theta \geq 0$ and $\sigma > 0$ are unknown parameters and where r is a positive integer. The goal is to construct a $1-\alpha$ CI for the parameter of interest θ.

The joint probability mass function of $W = w$ and $X = x$ is

$$f(w, x|\theta, \sigma) = \frac{1}{2^{\frac{r+1}{2}}\sigma^{r+1}\sqrt{\pi}\Gamma(r/2)} w^{\frac{r}{2}-1} \exp\left[-\frac{(x-\theta)^2 + w}{2\sigma^2}\right].$$

Let $s^2 = w/r$ and $H_r(\cdot)$ be the cumulative distribution function (cdf) of the t distribution with r degrees of freedom. Suppose that θ and σ are given the improper prior $f(\theta, \sigma) = \frac{1}{\sigma}$ for $0 < \sigma < \infty$ and $0 \leq \theta < \infty$. Then, the marginal density of X and W is

$$f(w, x) = \int_0^\infty \int_0^\infty \frac{f(w, x|\theta, \sigma)}{\sigma} d\sigma d\theta = \frac{1}{2w} H_r(t),$$

where $t = x/s$ and $s^2 = w/r$. Hence, the posterior density of θ is

$$g(\theta|w, x) = \int_0^\infty \frac{f(w, x|\theta, \sigma)}{f(w, x)\sigma} d\sigma = \frac{1}{sH_r(t)} h_r\left(\frac{\theta - x}{s}\right),$$

where $h_r(\cdot)$ is the density of $H_r(\cdot)$.

A level $1-\alpha$ Bayesian credible interval for θ requires two functions $l = l(w, x)$ and $u = u(w, x)$, for which

$$1 - \alpha = \int_l^u g(\theta|w, x)\, d\theta,$$

and to minimize the length of the interval, l and u must satisfy

$$[l, u] = \{\theta : g(\theta|w, x) \geq c\},$$

for some constant c. Hence, $l = max\{0, x - bs\}$ and $u = x + bs$, where

$$b = \max\left\{H_r^{-1}\left[1 - \alpha H_r(t)\right], H_r^{-1}\left[\frac{1}{2} + \frac{1}{2}(1-\alpha)H_r(t)\right]\right\},$$

and $t = x/s$.

2.2. Inference about the Poisson rate from a contaminated observed count

Suppose we observe $X = B + S$ and W, where B, S, and W are mutually independent random variables such that $B \sim Poisson(\varepsilon)$, $S \sim Poisson(\lambda)$, and $W \sim Poisson(m\varepsilon)$. Here, $\lambda \geq 0$ and $\varepsilon \geq 0$ are unknown parameters, whereas $m > 0$ is a known constant. The goal is to infer λ.

The joint probability mass function of $W = w$ and $X = x$ is

$$p\left(w, x|\varepsilon, \lambda\right) = \sum_{k=0}^{x} \frac{1}{w!k!(x-k)!} m^w \varepsilon^{w+k} \lambda^{x-k} e^{-\lambda-(m+1)\varepsilon}.$$

Zhang and Woodroofe [8] proposed a Bayesian solution for inferring λ. Suppose that λ and ε are given an improper prior density, e.g., $\varepsilon^{a-1}e^{-b\varepsilon}$, for $0 \leq \varepsilon, \lambda < \infty$, where $a > 0$, $b \geq 0$. Then the marginal probability mass function of W and X is

$$p(w, x) = \int_0^\infty \int_0^\infty p\left(w, x|\varepsilon, \lambda\right) \varepsilon^{a-1} e^{-b\varepsilon} d\varepsilon d\lambda = \sum_{k=0}^{x} \frac{\Gamma(a+w+k)}{w!k!} \frac{m^w}{(b+m+1)^{a+w+k}}.$$

Using Bayes' theorem, the posterior density and distribution function of λ are

$$g\left(\lambda|w, x\right) = \frac{1}{p(w, x)} \sum_{k=0}^{x} \frac{\Gamma(a+w+k)}{w!k!(x-k)!} \frac{m^w}{(b+m+1)^{a+w+k}} \varepsilon^{-\lambda} \lambda^{x-k}$$

and

$$G\left(\lambda|w, x\right) = \frac{1}{p(w, x)} \sum_{k=0}^{x} \frac{\Gamma(a+w+k)}{w!k!} \frac{m^w}{(b+m+1)^{a+w+k}} H_{x-k+1}(\lambda),$$

where $H_j(\cdot)$ denotes the gamma distribution function with shape parameter j and a unit scale parameter. Hence, a Bayesian credible interval consists of two functions, $l = l(w, x)$ and $u = u(w, x)$, for which

$$1 - \alpha = G\left(u|w, x\right) - G\left(l|w, x\right),$$

$$[l, u] = \left\{\lambda : g\left(\lambda|w, x\right) \geq c\right\},$$

for some $c = c(w, x)$.

3. New IM-based confidence intervals

In this section, we first briefly introduce the IM framework aimed at establishing valid probabilistic inference and then propose new IM-based CIs for the parameter of interest λ.

3.1. IM framework for valid probabilistic inference

Let X denote the observable sample data. The sampling model is defined as a probability distribution $P_{X|\theta}$ over the sample space $\mathbf{X}$, where $\theta \in \mathbf{\Theta}$ is the indexing parameter. For a given θ, this sampling model for X is induced via an auxiliary variable U. Specifically, let $\mathbf{U}$ be an auxiliary space endowed with a probability measure P_U and the sample model $P_{X|\theta}$ is determined via the following procedure:

$$\text{sample } U \sim P_U \text{ and set } X = \varphi(\theta, U),$$

where $\varphi : \mathbf{\Theta} \times \mathbf{U} \to \mathbf{X}$ denotes an appropriate mapping. Furthermore, if $X = x$ is observed, then we know that $x = \varphi(\theta, u^*)$, where u^* is some unobserved realization of U. Clearly, knowing u^* is equivalent to knowing θ. Hence, the IM approach attempts to accurately predict the value of u^* before conditioning on $X = x$.

Definition 3.1 (Martin and Liu [14])**.** A PRS S is valid for predicting the unobserved auxiliary variable if $Q_S(U) = P_S\{U \notin S\}$, as a function of $U \sim P_U$, is stochastically no larger than $Unif(0,1)$, that is, for each $\alpha \in (0,1)$, $P_U\{Q_S(U) \geq 1-\alpha\} \leq \alpha$. If "$\leq \alpha$" can be replaced by "$= \alpha$", then S is efficient.

The IM consists of three steps:

Association step: From an appropriate mapping $\varphi : X = \varphi(\theta, U)$, the IM establishes an association between the parameter θ and each possible pair (X, U). This association yields a collection of candidate value sets defined as $\Theta_X(U) = \{\theta : X = \varphi(\theta, U)\}$.

Prediction step: Given observed data $X = x$, let θ^* denote the true value of θ. There exists a corresponding u^* such that $x = \varphi(\theta^*, u^*)$. Furthermore, a valid PRS S is employed to predict this true u^*. The validity condition ensures that S will hit u^* with a large probability.

Combination step: The association step and prediction step are combined to obtain a final random set of θ, that is, $\Theta_x(S) = \bigcup_{u\in S} \Theta_x(u)$. Then, for any assertion A regarding the parameter of interest θ, two probability measures are computed to quantify the evidence in x supporting A: the belief function $bel_x(A) = P_S\{\Theta_x(S) \subseteq A\}$ and the plausibility function $pl_x(A) = P_S\{\Theta_x(S) \nsubseteq A^c\}$.

Clearly, $bel_x(A)$ and $pl_x(A)$ represent the minimum and maximum probabilities supporting the truth of assertion A, respectively. In practical applications, reporting the plausibility function is more convenient, as it can be easily utilized to construct frequentist statistical procedures. To test assertion $A = \{\theta : \theta = \theta_0\}$, the null hypothesis $H_0 : \theta = \theta_0$ is rejected at a significance level α if $pl_x(A) \leq \alpha$. Additionally, this plausibility function yields a two-sided $1-\alpha$ IM CI $\{\theta : pl_x(A) > \alpha\}$.

Definition 3.2 (Martin and Liu [14])**.** Suppose that $X \sim P_{X|\theta}$ and let A be an assertion of interest; the IM with the plausibility function $pl_X(A)$ is valid for assertion A if, for each $\alpha \in (0,1)$,

$$\sup_{\theta\in A} P_{X|\theta}\{pl_X(A) \leq \alpha\} \leq \alpha.$$

The IM is valid if it is valid for all A.

Theorem 3.1. Suppose that the PRS S is valid and that $Q_x(S) \neq \varnothing$ with a P_S probability of 1 for all x. Then, the IM CI $\{\theta : pl_x(A) > \alpha\}$ can guarantee the nominal coverage probability. Moreover, if "$\leq \alpha$" can be replaced by "$= \alpha$", then the IM interval controls the coverage probability exactly at the confidence level $1-\alpha$.

3.2. Inference for the constrained normal case

Suppose that independent random variables $X \sim N(\theta, \sigma^2)$ and $W \sim \sigma^2\chi^2_r$ are observed, where $\theta \in C_1 = \{\theta : \theta \geq 0\}$ and $\sigma > 0$ are unknown parameters and r is a positive integer. Let $Z \sim N(0,1)$, $Q \sim \chi^2_r$, where Z and Q are independent unobservable but predictable auxiliary variables. The association linking X, θ, and an auxiliary variable Z may be written as follows:

$$X = \theta + \sigma Z,\ Z \sim N(0,1). \tag{3.1}$$

Similarly, the association for W, given σ^2, may be written as

$$W = \sigma^2 Q, Q \sim \chi^2_r. \tag{3.2}$$

From (3.2), $\sigma^2 = W/Q$. We then plug $\sigma^2 = W/Q$ into (3.1), that is,

$$X = \theta + \sqrt{W}\frac{Z}{\sqrt{Q}},\ Z \sim N(0,1), Q \sim \chi^2_r\,. \tag{3.3}$$

Hence, the IM method has the following three-step procedures:

Association step: Let $F(\cdot)$ be the cdf of $Z/\sqrt{Q}$, which follows a t distribution with r degrees of freedom. From (3.3), the association step of the IM produces an initial candidate collection of θ,

$$\Theta_{X,W}(V) = \left\{\theta : X = \theta + \sqrt{W}F^{-1}(V)\right\},\ V \sim Unif(0,1). \tag{3.4}$$

Prediction step: We are interested in two-sided CIs, and the construction methods for one-sided CIs are analogous. Hence, for a singleton assertion $A = \{\theta : \theta = \theta_0\}$, the valid PRS of V proposed by Martin and Liu [14] is given by

$$S = \left\{V : |V - 0.5| \le \left|\widetilde{V} - 0.5\right|\right\},\ \widetilde{V} \sim Unif(0,1).$$

***Combination step*:** Combine $\Theta_{X,W}(V)$ and S to obtain a new candidate collection of θ, that is,

$$\Theta_{X,W}(S) = \bigcup_{V\in S}\Theta_{X,W}(V) = \left[X - \sqrt{W}F^{-1}\left(0.5 + \left|\widetilde{V} - 0.5\right|\right), X - \sqrt{W}F^{-1}\left(0.5 - \left|\widetilde{V} - 0.5\right|\right)\right].$$

Because θ is constrained to $C_1 = \{\theta : \theta \ge 0\}$, a conflict arises when $X - \sqrt{W}F^{-1}\left(0.5 + \left|\widetilde{V} - 0.5\right|\right) < 0$. To avoid the conflict case $\Theta_{X,W}(S) \bigcap C_1 = \varnothing$, one possible way is to enlarge the PRS $\Theta_{X,W}(S)$ for θ as follows:

$$\Theta'_{X,W}(S) = \begin{cases}\Theta_{X,W}(S) \cap C_1, & \text{if } \Theta_{X,W}(S) \cap C_1 \neq \varnothing,\\ \{0\}, & \text{if } \Theta_{X,W}(S) \cap C_1 = \varnothing.\end{cases}$$

Hence, the closed form of the final PRS $\Theta'_{X,W}(S)$ is

$$\Theta'_{X,W}(S) = \left[\max\left\{0, X - \sqrt{W}F^{-1}\left(0.5 + \left|\widetilde{V} - 0.5\right|\right)\right\}, \max\left\{0, X - \sqrt{W}F^{-1}\left(0.5 - \left|\widetilde{V} - 0.5\right|\right)\right\}\right].$$

For an assertion $A = \{\theta : \theta = \theta_0\}$, we calculate the plausibility function when $\theta_0 = 0$,

$$pl_{X,W}(A) = P_S\left\{\Theta'_{X,W}(S) \cap \{0\} \neq \varnothing\right\} = \begin{cases}2\left[1 - F\left(\dfrac{X}{\sqrt{W}}\right)\right], & \dfrac{1}{2} < F\left(\dfrac{X}{\sqrt{W}}\right) \le 1,\\ 1, & \text{otherwise,}\end{cases}$$

and for $\theta_0 > 0$,

$$pl_{X,W}(A) = P_S\left\{\Theta'_{X,W}(S) \cap A \neq \varnothing\right\} = \begin{cases}2F\left(\dfrac{X-\theta_0}{\sqrt{W}}\right), & F\left(\dfrac{X-\theta_0}{\sqrt{W}}\right) < \dfrac{1}{2},\\ 2\left[1 - F\left(\dfrac{X-\theta_0}{\sqrt{W}}\right)\right], & F\left(\dfrac{X-\theta_0}{\sqrt{W}}\right) \ge \dfrac{1}{2}.\end{cases}$$

Given the significance level α, we confirm that the assertion $A = \{\theta : \theta = \theta_0\}$ is wrong if $pl_{X,W}(A) \le \alpha$. Moreover, the IM two-sided $1 - \alpha$ CI for θ is

$$\left\{\theta : pl_{X,W}(A) > \alpha\right\} = \left[\max\left\{0, X - \sqrt{W}F^{-1}\left(1 - \frac{\alpha}{2}\right)\right\}, \max\left\{0, X - \sqrt{W}F^{-1}\left(\frac{\alpha}{2}\right)\right\}\right].$$

Theorem 3.2. The plausibility function $pl_{X,W}(A)$ of the IM method is valid for assertion $A = \{\theta : \theta = \theta_0\}$ if, for each $\alpha \in (0, 1)$,

$$\sup_{\theta \in A} P_{(X,W)|\theta,\sigma^2} \left\{ pl_{X,W}(A) \leq \alpha \right\} \leq \alpha.$$

Hence, the resulting IM CI can guarantee the nominal coverage probability $1 - \alpha$.

3.3. Inference in the Poisson case

Suppose that we observe $X = B + S$ and W, where B, S, and W are independent random variables for which $B \sim Poisson(\varepsilon)$, $S \sim Poisson(\lambda)$, and $W \sim Poisson(m\varepsilon)$, where $\lambda \in C_2 = \{\lambda : \lambda \geq 0\}$ and $\varepsilon \geq 0$ are unknown parameters and $m > 0$ is known. Let $\theta = \varepsilon + \lambda$; the goal is to infer λ. In this section, we propose two new IM-based CIs for λ.

3.3.1. IM confidence interval

Let $F_\theta(\cdot)$ and $F_{m\varepsilon}(\cdot)$ be the cdfs of $X \sim Poisson(\theta)$ and $W \sim Poisson(m\varepsilon)$, respectively. Owing to the discreteness of the Poisson distribution, Martin and Liu [14] gave an association linking X, θ, and an auxiliary variable U as follows:

$$X = \min\{k : U < F_\theta(k)\},\ U \sim \text{Unif}(0, 1),$$

which is equivalent to

$$F_\theta(X - 1) \leq U < F_\theta(X),\ U \sim Unif(0, 1). \tag{3.5}$$

Similarly, given $m\varepsilon$, the association model for W, can be written as follows:

$$F_{m\varepsilon}(W - 1) \leq \widetilde{U} < F_{m\varepsilon}(W), \widetilde{U} \sim Unif(0, 1), \tag{3.6}$$

where U and $\widetilde{U}$ are mutually independent. Let $G_{a,b}(\cdot)$ be the cdf of the gamma distribution with shape parameter a and scale parameter b , and let $G^{-1}_{a,b}(\cdot)$ be the inverse function of $G_{a,b}(\cdot)$. From the well-known relation between the Poisson and gamma distributions, $G_{X,1}(\theta) = 1 - F_\theta(X - 1)$, we can rewrite the joint associations (3.5) and (3.6) as follows:

$$G^{-1}_{X,1}(1 - U) \leq \theta < G^{-1}_{X+1,1}(1 - U), U \sim Unif(0, 1), \tag{3.7}$$

$$G^{-1}_{W,1}\left(1 - \widetilde{U}\right) \leq m\varepsilon < G^{-1}_{W+1,1}\left(1 - \widetilde{U}\right), \widetilde{U} \sim Unif(0, 1). \tag{3.8}$$

Hence, by (3.7) and (3.8), the initial association for $\lambda = \theta - \varepsilon$ is

$$G^{-1}_{X,1}(1 - U) - \frac{G^{-1}_{W+1,1}\left(1 - \widetilde{U}\right)}{m} < \lambda < G^{-1}_{X+1,1}(1 - U) - \frac{G^{-1}_{W,1}\left(1 - \widetilde{U}\right)}{m}. \tag{3.9}$$

Lemma 3.1. Let $X, W \in \{0, 1, \ldots\}$, $m > 0$, and let the auxiliary variables U and $\widetilde{U}$ be independent and identically distributed with $U, \widetilde{U} \sim \text{Unif}(0, 1)$. Define

$$L_{X,W}(U, \widetilde{U}) = G^{-1}_{X,1}(1 - U) - m^{-1}G^{-1}_{W+1,1}(1 - \widetilde{U}),$$

$$R_{X,W}(U,\widetilde{U}) = G^{-1}_{X+1,1}(1-U) - m^{-1}G^{-1}_{W,1}(1-\widetilde{U}).$$

Let $K_{X,W+1}(\cdot)$ and $K_{X+1,W}(\cdot)$ be the distribution functions of $L_{X,W}(U,\widetilde{U})$ and $R_{X,W}(U,\widetilde{U})$, respectively, and let $K^{-1}_{X,W+1}(\cdot)$ and $K^{-1}_{X+1,W}(\cdot)$ be their corresponding inverse functions. Then, for any $\lambda \geq 0$ and $V \sim \text{Unif}(0,1)$, we have

$$P_{U,\widetilde{U}}\left\{L_{X,W}(U,\widetilde{U}) < \lambda < R_{X,W}(U,\widetilde{U})\right\} = P_V\left\{K^{-1}_{X,W+1}(V) < \lambda < K^{-1}_{X+1,W}(V)\right\}.$$

The IM method has the following three-step procedures:

Association step: Using Lemma 3.1, the association step of the IM produces an initial candidate collection of λ:

$$\Theta_{X,W}(V) = \left\{\lambda : K^{-1}_{X,W+1}(V) < \lambda < K^{-1}_{X+1,W}(V)\right\},\ V \sim Unif(0,1). \tag{3.10}$$

Prediction step: For a singleton assertion $A = \{\lambda : \lambda = \lambda_0\}$, the valid PRS of V proposed by Martin and Liu [14] is given by

$$S = \left\{V : |V - 0.5| \leq \left|\widetilde{V} - 0.5\right|\right\},\ \widetilde{V} \sim Unif(0,1).$$

Combination step: $\Theta_{X,W}(V)$ and S are combined to obtain a random set $\Theta_{X,W}(S) = \bigcup_{V\in S}\Theta_{X,W}(V)$ for λ. Note that λ has a constraint $\lambda \in C_2$. Similar to the constrained normal case, to avoid the conflict case $\Theta_{X,W}(S)\bigcap C_2 = \varnothing$, we choose to enlarge the PRS $\Theta_{X,W}(S)$ for λ as follows:

$$\Theta'_{X,W}(S) = \begin{cases} \Theta_{X,W}(S)\cap C_2, & \text{if } \Theta_{X,W}(S)\cap C_2 \neq \varnothing, \\ \{0\}, & \text{if } \Theta_{X,W}(S)\cap C_2 = \varnothing. \end{cases}$$

Hence, the closed form of the final PRS $\Theta'_{X,W}(S)$ is

$$\Theta'_{X,W}(S) = \left[\max\left\{0, K^{-1}_{X,W+1}\left(0.5 - \left|\widetilde{V} - 0.5\right|\right)\right\}, \max\left\{0, K^{-1}_{X+1,W}\left(0.5 + \left|\widetilde{V} - 0.5\right|\right)\right\}\right].$$

Moreover, we calculate the plausibility function for an assertion $A = \{\lambda : \lambda = \lambda_0\}$ as follows:

$$pl_{X,W}(A) = P_S\left\{\Theta'_{X,W}(S)\cap\{0\} \neq \varnothing\right\} = \begin{cases} 2K_{X,W+1}(0), & K_{X,W+1}(0) < \frac{1}{2}, \\ 1, & \text{otherwise}, \end{cases}$$

when $\lambda_0 = 0$, and

$$pl_{X,W}(A) = P_S\left\{\Theta'_{X,W}(S)\cap A \neq \varnothing\right\} = \begin{cases} 2K_{X,W+1}(\lambda_0), & K_{X,W+1}(\lambda_0) < \frac{1}{2}, \\ 2[1 - K_{X+1,W}(\lambda_0)], & K_{X+1,W}(\lambda_0) > \frac{1}{2}, \\ 1, & \text{otherwise}, \end{cases}$$

when $\lambda_0 > 0$.

Theorem 3.3. According to Theorem 3.1, the plausibility function $pl_{X,W}(A)$ of the IM method is valid for assertion A if, for each $\alpha \in (0,1)$,

$$\sup_{\lambda\in A} P_{(X,W)|\lambda,\varepsilon}\left\{pl_{X,W}(A) \leq \alpha\right\} \leq \alpha.$$

For any $\alpha \in (0,1)$, if $pl_{X,W}(A) \leq \alpha$, then assertion A is wrong. Moreover, this plausibility function yields a two-sided IM $1-\alpha$ CI

$$CI_{IM} = \left\{\lambda : pl_{X,W}(A) > \alpha\right\} = \left[\max\left\{0, K^{-1}_{X,W+1}\left(\frac{\alpha}{2}\right)\right\}, \max\left\{0, K^{-1}_{X+1,W}\left(1-\frac{\alpha}{2}\right)\right\}\right].$$

Deriving closed-form expressions for $K_{X,W+1}(\cdot)$ and $K_{X+1,W}(\cdot)$ is challenging. We therefore approximate them using a Monte Carlo method, as described in Algorithm 1.

Algorithm 1: Calculating the IM CI $CI_{IM} = [\lambda_L, \lambda_U]$.

Input: Confidence level $1-\alpha$, scale parameter $m > 0$, data X and W, number of simulations N

Output: IM interval $[\lambda_L, \lambda_U]$

Randomly sample U and $\widetilde{U}$ independently N times from $Unif(0,1)$;

for $i \leftarrow 1$ **to** N **do**

Calculate:

$$\lambda_{1,i} = \max\left\{0, G^{-1}_{X,1}(1-U_i) - \frac{G^{-1}_{W+1,1}(1-\widetilde{U}_i)}{m}\right\}$$

$$\lambda_{2,i} = \max\left\{0, G^{-1}_{X+1,1}(1-U_i) - \frac{G^{-1}_{W,1}(1-\widetilde{U}_i)}{m}\right\}$$

end

Approximate λ_L as the $\alpha/2$ quantile of the N realizations $\{\lambda_{1,i}\}_{i=1}^N$;

Approximate λ_U as the $1-\alpha/2$ quantile of the N realizations $\{\lambda_{2,i}\}_{i=1}^N$;

return $[\lambda_L, \lambda_U]$

Remark 3.1. Since the distribution functions of $K^{-1}_{X,W+1}$ and $K^{-1}_{X+1,W}$ lack explicit expressions, we use Algorithm 1 to simulate quantiles of λ to approximate λ_L and λ_U. Notably, this approximation method guarantees the nominal coverage exactly. We demonstrate this important property through simulation studies. Without loss of generality, we set $\varepsilon = 3.0$, $m = 20$, $1-\alpha = 0.90$, and $\lambda = 1.0, 2.0$. Table 1 reports the coverage probabilities and expected lengths of the IM intervals for different values of N. The IM intervals maintain the nominal coverage for all the simulated N, and the coverage probabilities and expected lengths show little variation. In fact, by the Glivenko–Cantelli theorem, as $N \to \infty$, the simulated quantiles converge to the theoretical quantiles. To obtain stable and accurate IM intervals, larger N yields more reliable results. For example, for $N = 10{,}000, 100{,}000, 1{,}000{,}000$, the Monte Carlo errors of the resulting IM intervals are on the order of 0.01, 0.00316, and 0.001. In this paper, considering both the computational time and the accuracy of interval estimation, we recommend $N = 10{,}000$.

Table 1. Coverage probability, expected length, and runtime of IM intervals under different values of λ and simulation size N.

λ	N	Coverage probability	Expected length	Runtime (s)
1.0	1,000	0.9434	6.0201	0.2799
1.0	5,000	0.9475	6.0032	1.2131
1.0	10,000	0.9428	6.0414	2.4021
1.0	50,000	0.9471	6.0203	11.5064
1.0	100,000	0.9518	6.0315	22.0841
2.0	1,000	0.9414	7.2047	0.2704
2.0	5,000	0.9368	7.1964	1.2233
2.0	10,000	0.9386	7.2532	2.4221
2.0	50,000	0.9379	7.2435	11.6211
2.0	100,000	0.9392	7.2077	22.3000

Note: The definitions of coverage probability and expected length can be found in Section 4.

3.3.2. Nonrandomized IM CI using a random weighting approach

In general, the association model (3.10) of the IM is an interval, which may result in a conservative CI. Certain adjustments are required to address this discreteness. Inspired by the random weighting idea (Lu et al. [15]), we consider a nonrandomized IM (NIM) approach to modify inequalities (3.5) and (3.6) to obtain equations so that we can improve the accuracy of the candidate value of λ. Specifically, let $F_\theta(\cdot)$ be the cdf of the Poisson distribution with parameter θ. For any fixed $\omega \in [0, 1]$ and a nonnegative integer $X \in [0, \infty)$, let

$$J_{X,\omega}(\theta) = \begin{cases} \omega I_0(\theta) + (1-\omega)F_\theta(X), & \text{if } X = 0, \\ \omega F_\theta(X-1) + (1-\omega)F_\theta(X), & \text{if } X \in (0, \infty), \end{cases}$$

where $I_D(\cdot)$ is the indicator function of set D. It is a strictly decreasing function of $\theta \in [0, \infty)$ and has a range of $[0, 1]$. Moreover, given X, for every ω, $u \in \text{Unif}(0, 1)$, there exists a unique solution

$$\theta = J^{-1}_{X,\omega}(u) = \sup\{\theta : J_{X,\omega}(\theta) \geq u\}. \tag{3.11}$$

Similarly, for any fixed $m > 0$, $\widetilde{\omega} \in [0, 1]$ and a nonnegative integer $W = w \in [0, +\infty)$, let

$$\widetilde{J}_{W,\widetilde{\omega}}(\varepsilon) = \begin{cases} \widetilde{\omega} I_0(m\varepsilon) + (1-\widetilde{\omega})\, F_{m\varepsilon}(W), & \text{if } W = 0, \\ \widetilde{\omega} F_{m\varepsilon}(W-1) + (1-\widetilde{\omega})\, F_{m\varepsilon}(W), & \text{if } W \in (0, \infty). \end{cases}$$

For every $\widetilde{\omega}, \widetilde{u} \in Unif(0, 1)$, there exists a unique solution

$$\varepsilon = \widetilde{J}^{-1}_{W,\widetilde{\omega}}(\widetilde{u}) = \sup\left\{\varepsilon : \widetilde{J}_{W,\widetilde{\omega}}(\varepsilon) \geq \widetilde{u}\right\}. \tag{3.12}$$

From (3.11) and (3.12), we have $\lambda = \theta - \varepsilon = J^{-1}_{X,\omega}(u) - \widetilde{J}^{-1}_{W,\widetilde{\omega}}(\widetilde{u})$. Let $H_{X,W}(\cdot)$ be its distribution function, and the NIM CI has the following three steps:

Association step: The new NIM method produces a new association for λ, that is,

$$\Theta_{X,W}(v) = \left\{\lambda : \lambda = H_{X,W}^{-1}(v)\right\}, \quad v \sim \text{Unif}(0,1).$$

Prediction step: For a singleton assertion $A = \{\lambda : \lambda = \lambda_0\}$, the valid PRS for v is

$$S = \{v : |v - 0.5| \leq |V - 0.5|\}, \quad V \sim \text{Unif}(0,1).$$

Combination step: To obtain an initial PRS for λ, $\Theta_{X,W}(v)$ and S are combined as follows:

$$\Theta_{X,W}(S) = \bigcup_{v \in S} \Theta_{X,W}(v) = \left[H_{X,W}^{-1}(0.5 - |V - 0.5|), H_{X,W}^{-1}(0.5 + |V - 0.5|)\right].$$

Since $\lambda \in C_2$, the final PRS for λ is obtained by enlarging $\Theta_{X,W}(S)$ to

$$\Theta'_{X,W}(S) = \left[\max\left\{0, H_{X,W}^{-1}(0.5 - |V - 0.5|)\right\}, \max\left\{0, H_{X,W}^{-1}(0.5 + |V - 0.5|)\right\}\right].$$

Then, the plausibility function of an assertion $A = \{\lambda : \lambda = \lambda_0\}$ is given by

$$pl_{X,W}(A) = P_S\left\{\Theta'_{X,W}(S) \cap \{0\} \neq \varnothing\right\} = \begin{cases} 2H_{X,W}(0), & H_{X,W}(0) < \frac{1}{2}, \\ 1, & \text{otherwise}, \end{cases}$$

when $\lambda_0 = 0$, and

$$pl_{X,W}(A) = P_S\left\{\Theta'_{X,W}(S) \cap A \neq \varnothing\right\} = \begin{cases} 2H_{X,W}(\lambda_0), & H_{X,W}(\lambda_0) < \frac{1}{2}, \\ 2\left[1 - H_{X,W}(\lambda_0)\right], & H_{X,W}(\lambda_0) \geq \frac{1}{2}, \end{cases}$$

when $\lambda_0 > 0$.

Theorem 3.4. Let $S \sim P_S$ be a valid PRS for $v \sim \text{Unif}(0,1)$; that is, $P_S\{v \in S\} \geq_{st} \text{Unif}(0,1)$, where "$\geq_{st}$" means "stochastically no smaller than". If $H_{X,W}(\lambda) \sim \text{Unif}(0,1)$ for $(X,W) \sim P_{(X,W)|\lambda}$ for all λ, then the NIM method is valid.

Given the significance level α, the resulting two-sided $1 - \alpha$ NIM CI for λ is

$$CI_{NIM} = \left\{\lambda : pl_{X,W}(A) > \alpha\right\} = \left[\max\left\{0, H_{X,W}^{-1}\left(\frac{\alpha}{2}\right)\right\}, \max\left\{0, H_{X,W}^{-1}\left(1 - \frac{\alpha}{2}\right)\right\}\right].$$

Since it is difficult to obtain the closed form of $H_{X,W}(\lambda)$, we suggest approximating it via a Monte Carlo method as follows:

Algorithm 2: Calculating the NIM CI $CI_{NIM} = [\lambda_L, \lambda_U]$.

Input: Confidence level $1-\alpha$, $m > 0$, data X, W, number of simulations N

Output: NIM interval $[\lambda_L, \lambda_U]$

for $i \leftarrow 1$ **to** N **do**

1. Sample $\omega, \widetilde{\omega}, u, \widetilde{u} \overset{iid}{\sim} \text{Unif}(0,1)$;
2. Solve for θ using:

$$u = \begin{cases} \omega I_0(\theta) + (1-\omega)F_\theta(X), & \text{if } X = 0, \\ \omega F_\theta(X-1) + (1-\omega)F_\theta(X), & \text{if } X \in (0,\infty); \end{cases}$$

3. Solve for ε using:

$$\widetilde{u} = \begin{cases} \widetilde{\omega} I_0(m\varepsilon) + (1-\widetilde{\omega})F_{m\varepsilon}(W), & \text{if } W = 0, \\ \widetilde{\omega} F_{m\varepsilon}(W-1) + (1-\widetilde{\omega})F_{m\varepsilon}(W), & \text{if } W \in (0,\infty); \end{cases}$$

4. Calculate $\lambda^{(i)} = \max\{0, \theta - \varepsilon\}$;

end

Calculate λ_L as the $\alpha/2$ quantile of the realizations $\{\lambda^{(i)}\}_{i=1}^N$;

Calculate λ_U as the $1-\alpha/2$ quantile of the realizations $\{\lambda^{(i)}\}_{i=1}^N$;

return $[\lambda_L, \lambda_U]$

Remark 3.2. According to Theorem 3.4, the NIM method is valid if the Monte Carlo approximation $H_{X,W}(\lambda)$ exactly follows a uniform distribution on $(0,1)$. Although this condition is not always satisfied, the simulation study in Section 4.2 shows that the approximation of $H_{X,W}(\lambda)$ to a uniform distribution on $(0,1)$ can provide a theoretical justification for the superior coverage performance of the NIM intervals. Specifically, for a given background parameter ε, the approximation of $H_{X,W}(\lambda)$ to Unif$(0,1)$ is better when the background rate parameter m is small, but deteriorates when m is large. In fact, when m is small, the mean $m\varepsilon$ of W is also small for fixed ε, so the values of W are typically concentrated on a few small integers, indicating strong discreteness. In this case, the random weighting mechanism of the NIM method can effectively mitigate the discreteness of the data, making the distribution function $H_{X,W}(\lambda)$ of λ very close to Unif$(0,1)$. Conversely, when m is large, the mean of W is large, and its distribution is approximately continuous. Hence, discreteness is no longer the main issue. As a result, the extra randomness introduced by random weighting becomes an unnecessary perturbation, causing $H_{X,W}(\lambda)$ to deviate from uniformity to a greater extent than in the small-m case.

Remark 3.3. When evaluating the coverage performance of the NIM CIs, we first generate M observed data pairs $\{(X_i, W_i)\}_{i=1}^M$. For each (X_i, W_i), we then generate N random vectors $(\omega, \tilde{\omega}, u, \tilde{u})$ to compute the $1-\alpha$ NIM CI, which requires solving N nonlinear equations for each (X_i, W_i). According to our simulation study, although the `uniroot` function in R software is suitable for solving a single nonlinear equation, its serial computing strategy leads to a prohibitive computational time when equations are solved in batches. To address this issue, we instead employ the bisection method in Python and leverage GPUs (via PyTorch tensors) for large-scale parallel computing, allowing all the equations to be processed simultaneously. For example, when $m = 20$, $\lambda = 1$, $\varepsilon = 3$, $M = N = 10{,}000$, and $1-\alpha = 0.90$, the coverage probability computed by R is 0.9106 with a runtime of approximately 11

hours, whereas Python yields 0.9083 in approximately 15 seconds. Thus, the GPU-based parallel strategy is not only accurate but also drastically reduces the computation time. We therefore adopt this strategy in the simulation studies reported in Section 4. All R and Python codes used in this article are available from the corresponding author upon reasonable request.

Remark 3.4. In terms of GPU computation, we compute the NIM intervals using batched tensor operations. Specifically, we first generate M simulated observed pairs $\{(X_i, W_i)\}_{i=1}^{M}$ on the GPU, and then generate N random vectors $(\omega, \tilde{\omega}, u, \tilde{u})$ for each observed pair (X_i, W_i), forming an $M \times N$ tensor structure. Next, we employ a vectorized bisection method to solve the nonlinear equations for θ and ε in Algorithm 2 in parallel on the GPU. Finally, we compute the NIM interval for λ corresponding to each observed pair. Although parallel tensor computing on the GPU improves computational efficiency, the choice of the simulation size N affects both the accuracy of the interval estimates and the runtime. We conduct a sensitivity analysis for N. Without loss of generality, we fix $m = 20$, $\lambda = 1, 2$, $\varepsilon = 3$, $M = 10{,}000$, and $1 - \alpha = 0.90$. Let $N = 1{,}000, 5{,}000, 10{,}000, 50{,}000, 100{,}000$. Table 2 reports the coverage probabilities, expected lengths, and runtimes of the NIM intervals for different values of N. It is evident that the runtime of the NIM intervals increases approximately linearly with N. Moreover, when $N \geq 10{,}000$, the coverage probability and expected length of the NIM intervals stabilize. In fact, according to the central limit theorem, the simulation error of the NIM intervals decays at the rate $O(N^{-1/2})$. For example, when $N = 10{,}000$, we are 95% confident that the absolute error between the simulated and true values of the NIM intervals does not exceed 0.01. Therefore, to balance the numerical accuracy, runtime, and computational cost for future improvements in the NIM method, we recommend using $N = 10{,}000$ in practical applications.

Table 2. Coverage probability, expected length, and runtime of NIM intervals under different values of λ and simulation size N.

λ	N	Coverage probability	Expected length	Runtime (s)
1.0	1,000	0.9087	5.3656	1.3275
1.0	5,000	0.9104	5.3860	6.2577
1.0	10,000	0.9045	5.3815	12.3901
1.0	50,000	0.9080	5.4089	61.1492
1.0	100,000	0.9021	5.3657	121.6561
2.0	1,000	0.9057	6.5228	1.3370
2.0	5,000	0.9058	6.5947	6.3637
2.0	10,000	0.9061	6.5855	12.5732
2.0	50,000	0.9049	6.5593	62.1051
2.0	100,000	0.9029	6.5371	123.8227

Note: The simulation results were obtained using PyTorch 2.4.1+cu118 and Python 3.12.12. The definitions of coverage probability and expected length can be found in Section 4.

4. Simulation study

4.1. Constrained normal case

For the constrained normal mean θ, we have proven that the IM CI has the exact coverage probability. There is no need to perform simulations to demonstrate its validity. However, the complex expressions of Bayesian credible intervals make it difficult to theoretically compare the frequentist coverage performance between Bayesian and IM intervals. In this section, we perform Monte Carlo simulations to examine the frequentist performance of the two CIs. We mainly report the empirical coverage probability and expected length. These two measurements are commonly used to evaluate the reliability and precision of an interval. With respect to the setting of parameter values, we refer to Zhang and Woodroofe [8] and Leaf and Liu [9]. We set $\sigma^2 = 1$, $\theta = 0.0\ (0.1)\ 4.0$, $1-\alpha = 0.90,\ 0.95$, and $r = 5,\ 10,\ 20,\ 50$. For every given $(\theta, \sigma^2, r, \alpha)$, we first independently resample $X \sim N(\theta, \sigma^2)$ and $W \sim \sigma^2\chi^2_r$ 10,000 times, calculate the two CIs of θ accordingly, and compute the corresponding frequencies that cover θ. We regard the coverage frequency as the coverage probability. According to the central limit theorem, the coverage frequency of the nominal 95% confidence level tends to fall within the interval $(0.9457, 0.9543)$ for 10,000 experimental repetitions. For the expected length, by definition,

$$\begin{aligned}\text{Expected length} &= E_{\theta,\sigma^2}[\text{length}(CI(x,w))] \\ &= \frac{\sum_{i=1}^{M}\left[U\left(x,w;\theta,\sigma^2\right) - L\left(x,w;\theta,\sigma^2\right)\right]}{M},\end{aligned}$$

where $M = 10,000$ and $U\left(x,w;\theta,\sigma^2\right)$ and $L\left(x,w;\theta,\sigma^2\right)$ are the upper and lower limits of the interval for x and w, respectively. The simulation results are shown in Figures 1 and 2.

From Figures 1 and 2, the IM intervals maintain stable coverage probabilities for all the simulated values of θ, consistently remaining around the nominal confidence level lines, indicating that the IM intervals achieve exact coverage performance. In contrast, the coverage behavior of the Bayesian intervals is highly unstable. The coverage probability first increases to a peak, then decreases sharply to a level significantly below the nominal confidence level, and finally gradually approaches the nominal level. One possible reason for this undercoverage is that the improper uniform prior assigns insufficient weight near the boundary $\theta = 0$. Consequently, when the data X are moderate or large, the posterior distribution becomes overly concentrated away from zero. The IM method, however, does not rely on any prior and directly incorporates the constraint via the PRS, thereby maintaining exact coverage for all θ.

In terms of expected lengths, the expected lengths of the Bayesian and IM intervals monotonically increased with θ, and the expected length of the IM intervals was consistently longer than that of the Bayesian intervals. Notably, the coverage probability of the Bayesian interval is significantly lower than the nominal confidence level in many simulation scenarios, indicating that the IM method accurately controls the coverage probability by appropriately increasing the interval length. Moreover, the gap in expected lengths between the two types of intervals gradually narrowed as r increased. For example, when $r = 50$, the expected lengths of the two intervals were almost identical (Figure 2). Therefore, IM intervals achieve more reliable frequentist coverage at the cost of slightly longer interval lengths. Hence, the proposed IM CIs outperform the Bayesian intervals.

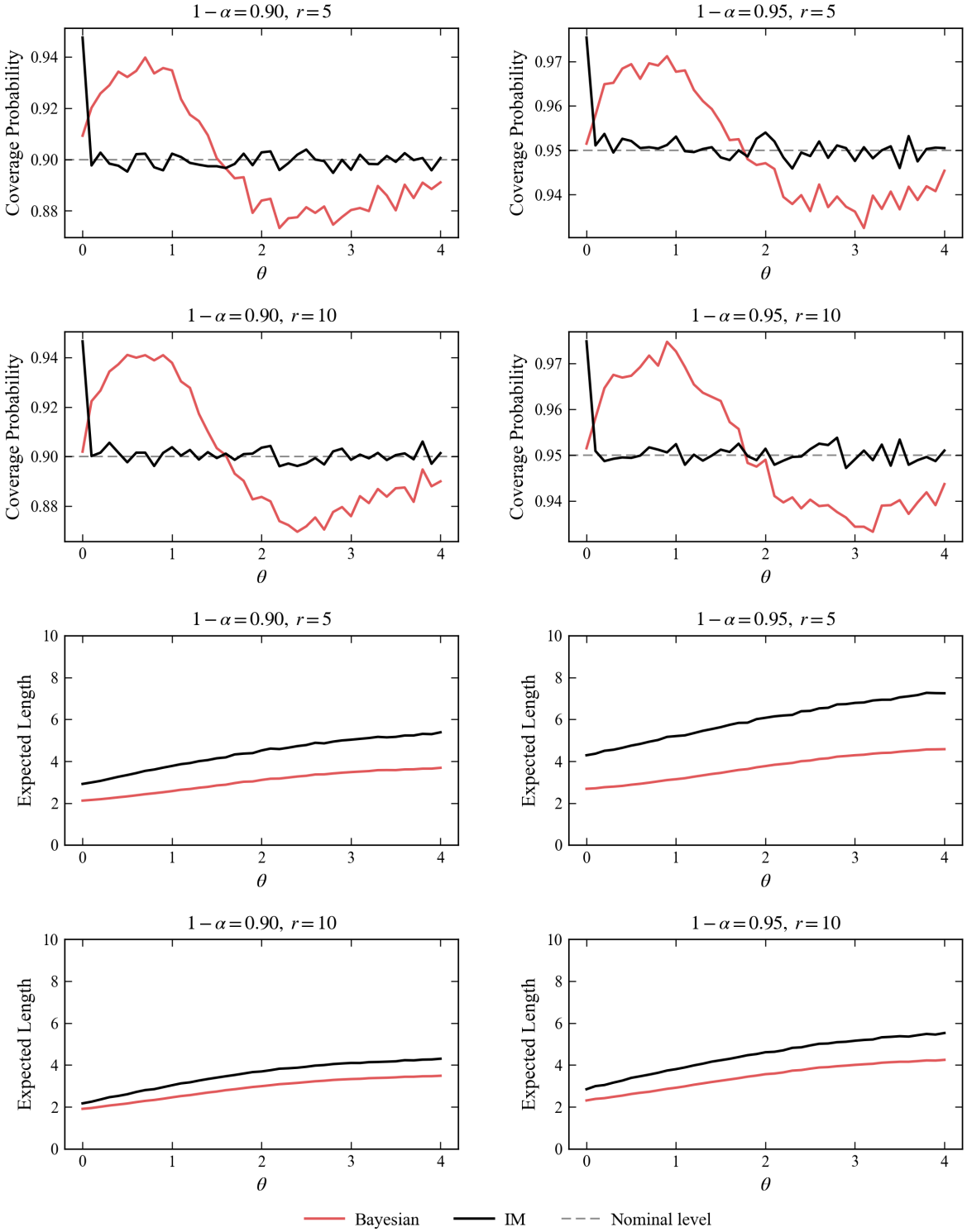

Figure 1. Coverage probabilities and expected lengths of the Bayesian and IM CIs for $\theta \in 0.0\ (0.1)\ 4.0$ when $r = 5,\ 10$ and $1 - \alpha = 0.90,\ 0.95$.

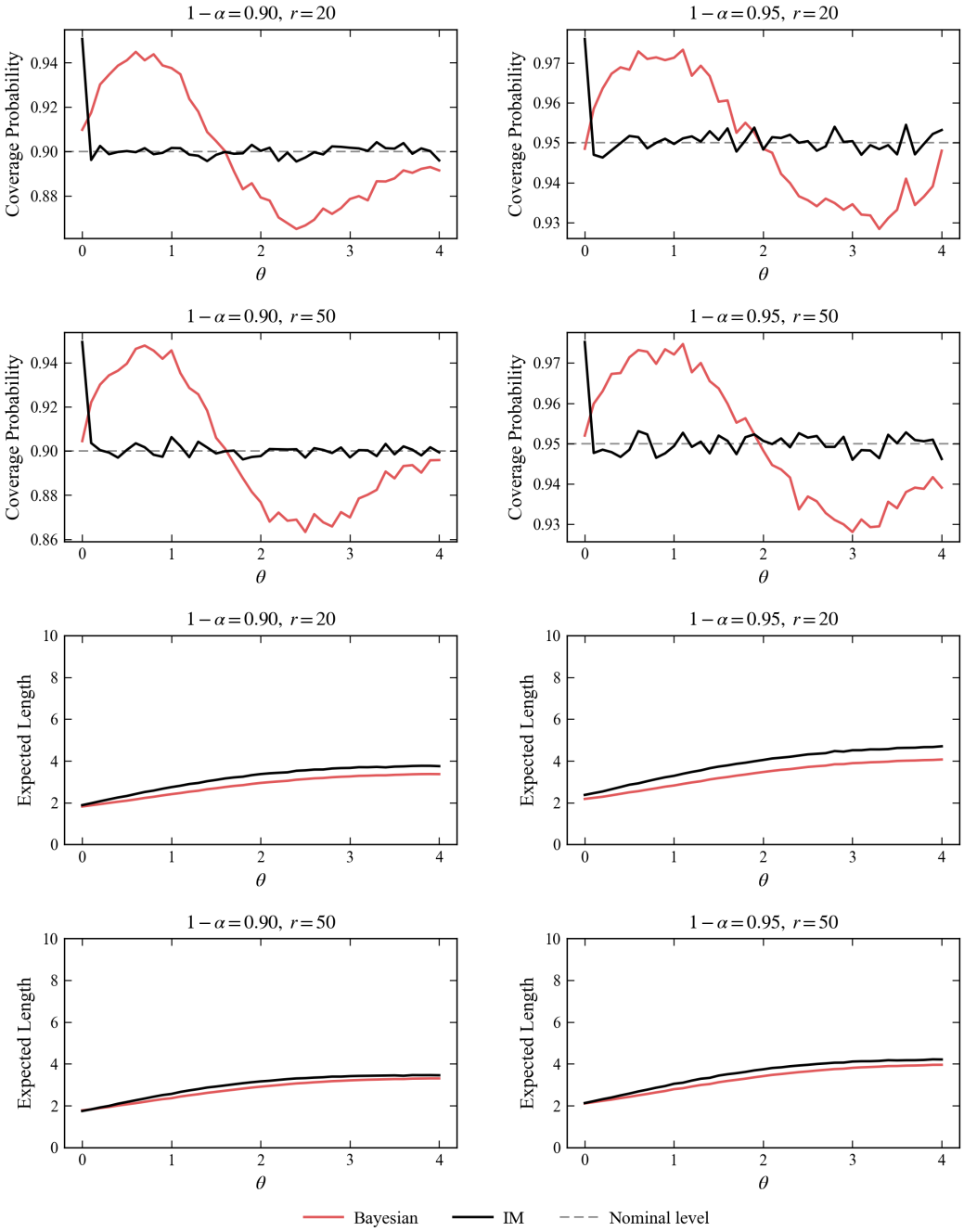

Figure 2. Coverage probabilities and expected lengths of the Bayesian and IM CIs for $\theta \in 0.0\ (0.1)\ 4.0$ when $r = 20,\ 50$ and $1 - \alpha = 0.90,\ 0.95$.

4.2. Constrained Poisson case

The lack of explicit expressions for Bayesian, IM, and NIM CIs precludes direct theoretical comparisons of their coverage performance. Therefore, Monte Carlo simulations are performed to evaluate the frequentist performance of these intervals, with coverage probability and expected length as the evaluation criteria. The parameter settings adopted from Zhang and Woodroofe [8] and Leaf and Liu [9] are as follows: $\varepsilon = 3.0$; $\lambda = 0.0\ (0.1)\ 10.0$; $1 - \alpha = 0.90,\ 0.95$; and $m = 20,\ 50,\ 100,\ 300$. The

simulation results are shown in Figures 3 and 4.

From Figures 3 and 4, as λ and m increase, the coverage probability of the Bayesian intervals becomes highly unstable, exhibiting an overall monotonically decreasing trend. This leads to a significant decrease below the preset confidence level in certain scenarios. In contrast, the proposed IM intervals consistently maintain stable coverage probabilities that are above the nominal confidence levels (0.90 and 0.95) for all values of λ and m, validating the effectiveness guaranteed by Theorem 3.3. Moreover, as an improved variant of the conservative IM interval, the NIM intervals exhibit smaller fluctuations in coverage probability and align most closely with the nominal confidence levels, demonstrating superior stability compared with both the Bayesian and IM intervals. This improvement is attributed to the random weighting technique, which corrects the discreteness of the Poisson distribution. By converting the inequality associations (3.5) and (3.6) into exact equations, the NIM method eliminates the artificial conservatism inherent in the IM interval while preserving frequentist validity.

In terms of the expected length, all three intervals increase monotonically with λ. The IM intervals are consistently slightly longer than the Bayesian interval, which is a reasonable trade-off for achieving stable coverage. Moreover, the NIM interval, as an improved version of the IM interval, further shortens the interval length to obtain more precise coverage performance. In addition, as m increases, the differences in the expected length among the three intervals remain nearly unchanged, indicating the robustness of the interval width. Furthermore, when $\lambda \leq 1.0$, the NIM interval is shorter; otherwise, the Bayesian interval is slightly shorter. Note that the IM interval does not rely on prior information. It provides stable and reliable interval estimation while overcoming the conservative and undercoverage issues of the Bayesian interval. Although the coverage probability of the NIM interval falls slightly below the nominal confidence level in some simulation scenarios, it optimizes the length of the IM interval to be close to or even shorter than that of the Bayesian interval. Moreover, its coverage probability is closest to the nominal level. Therefore, the new IM and NIM intervals outperform the Bayesian interval in both maintaining nominal coverage and achieving shorter interval lengths.

According to Theorem 3.4, for each $\alpha \in (0, 1)$, the NIM method is valid if the Monte Carlo approximation $H_{X,W}(\lambda)$ follows the uniform distribution in (0,1). However, our simulation studies show that this condition does not always hold, but the approximation $H_{X,W}(\lambda)$ can also provide some good properties. Without loss of generality, we construct different pairs $(\varepsilon,\ \lambda, m)$ to characterize the distribution function of $H_{X,W}(\lambda)$, where $\varepsilon = 3.0$; $\lambda = 0.0,\ 0.5,\ 1.0,\ 2.0,\ 3.0,\ 4.0$; and $m = 20,\ 50,\ 100,\ 300$. Specifically, for each λ and ε, we generate 10,000 Poisson random samples (X, W) and obtain a Monte Carlo estimate of the distribution function of $H_{X,W}(\lambda)$. Clearly, for small to moderate values of m, Figures 5 and 6 demonstrate that the approximate $H_{X,W}(\lambda)$ is sufficiently close to Unif(0,1). Moreover, for large m, Figures 7 and 8 show that the approximate $H_{X,W}(\lambda)$ has the following desirable property:

$$\begin{aligned}\sup_{\lambda\in A} P_{(X,W)|\lambda,\varepsilon}\{pl_{X,W}(A) \leq \alpha\} &= \sup_{\lambda\in A} P_{(X,W)|\lambda,\varepsilon}\left\{H_{X,W}(\lambda) \leq \frac{\alpha}{2} \cup H_{X,W}(\lambda) \geq 1 - \frac{\alpha}{2}\right\} \\ &\leq \sup_{\lambda\in A} P_{(X,W)|\lambda,\varepsilon}\left\{V \leq \frac{\alpha}{2} \cup V \geq 1 - \frac{\alpha}{2}\right\} = \alpha,\end{aligned}$$

where $V \sim \text{Unif}(0, 1)$. Hence, the approximate plausibility function of the NIM is valid.

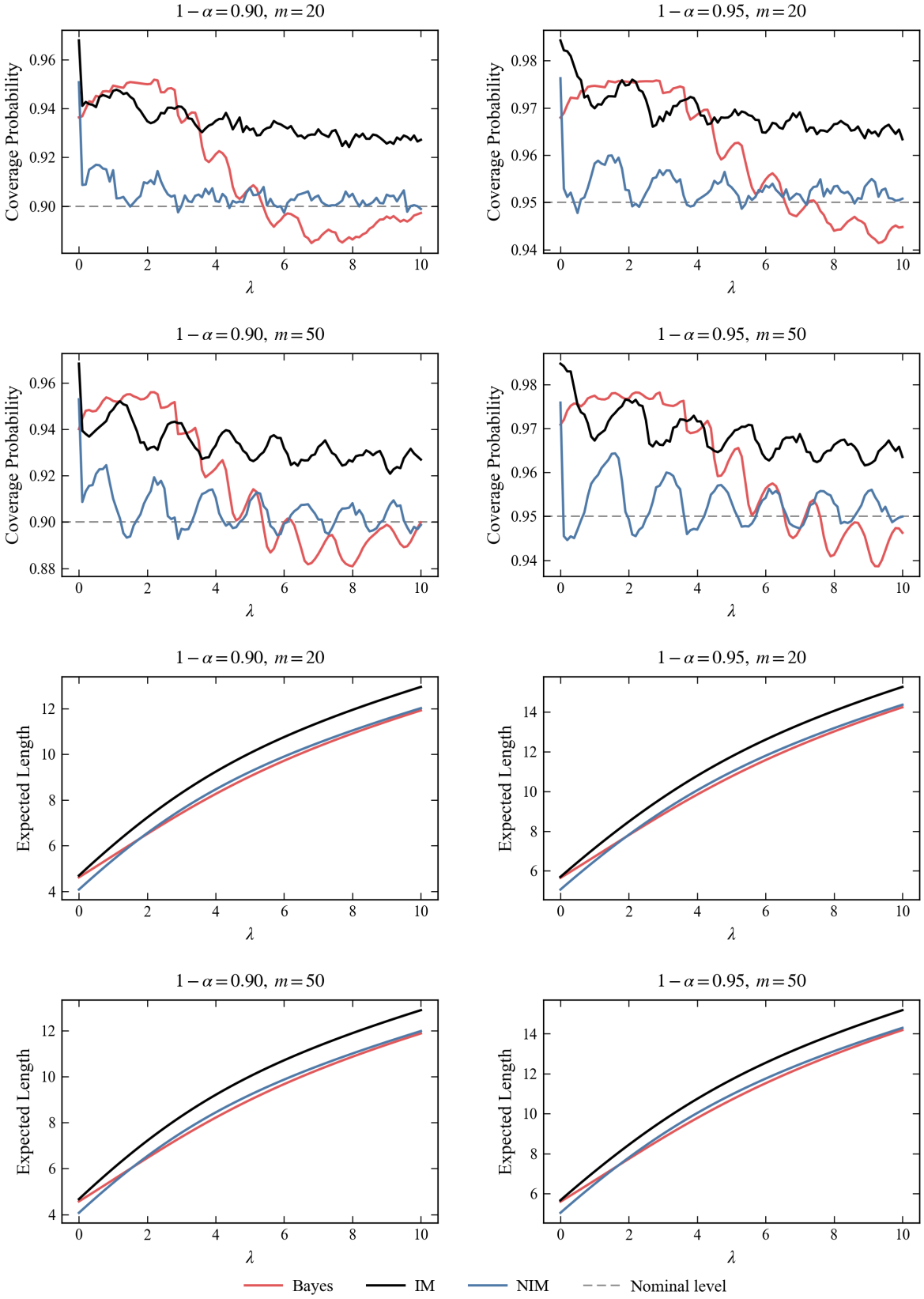

Figure 3. Coverage probabilities and expected lengths of the Bayesian, IM, and NIM CIs for $\lambda \in 0.0\ (0.1)\ 10.0$ when $\varepsilon = 3.0$, $m = 20,\ 50$ and $1 - \alpha = 0.90,\ 0.95$.

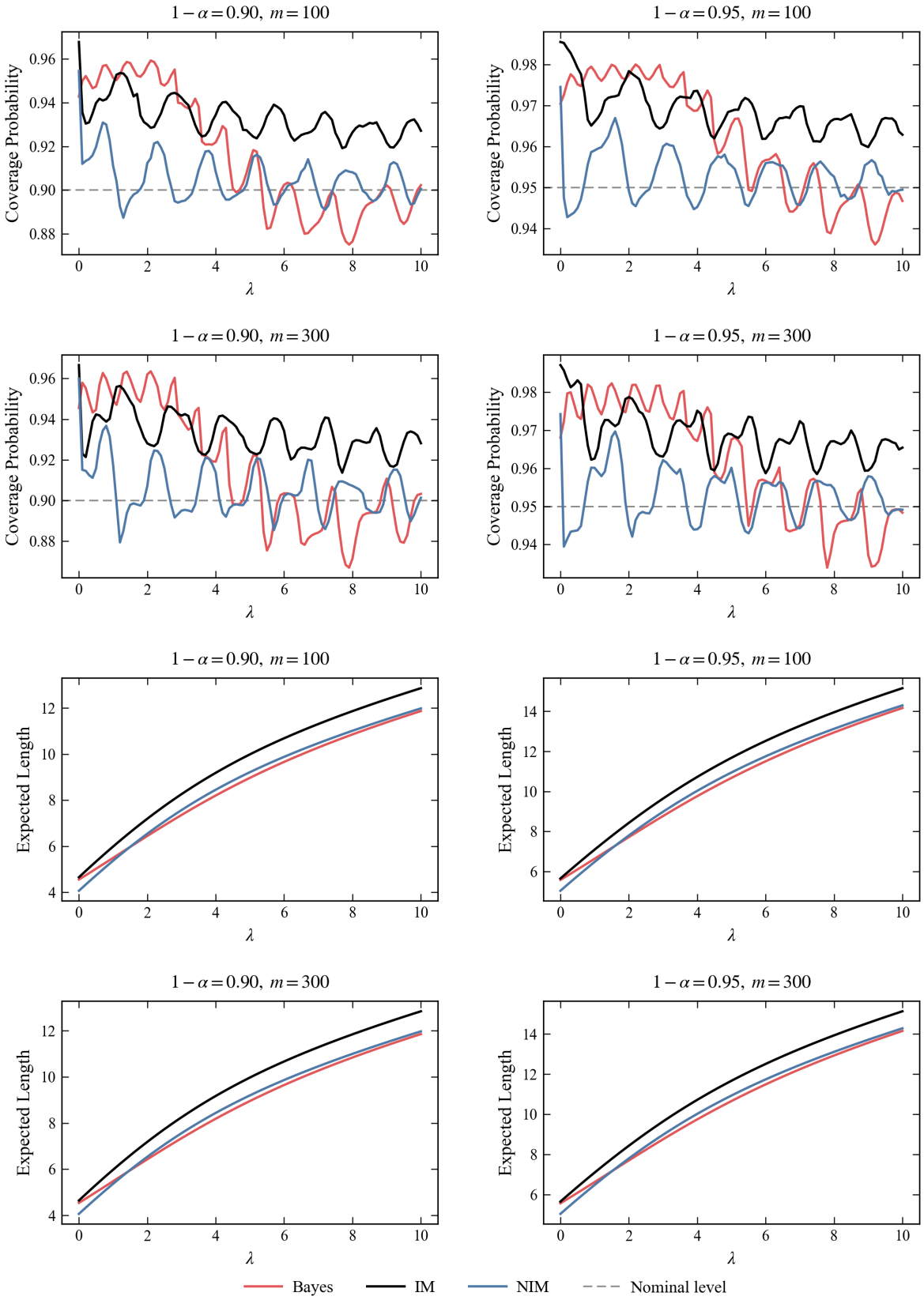

Figure 4. Coverage probabilities and expected lengths of the Bayesian, IM, and NIM CIs for $\lambda \in 0.0\ (0.1)\ 10.0$ when $\varepsilon = 3.0$, $m = 100,\ 300$ and $1 - \alpha = 0.90,\ 0.95$.

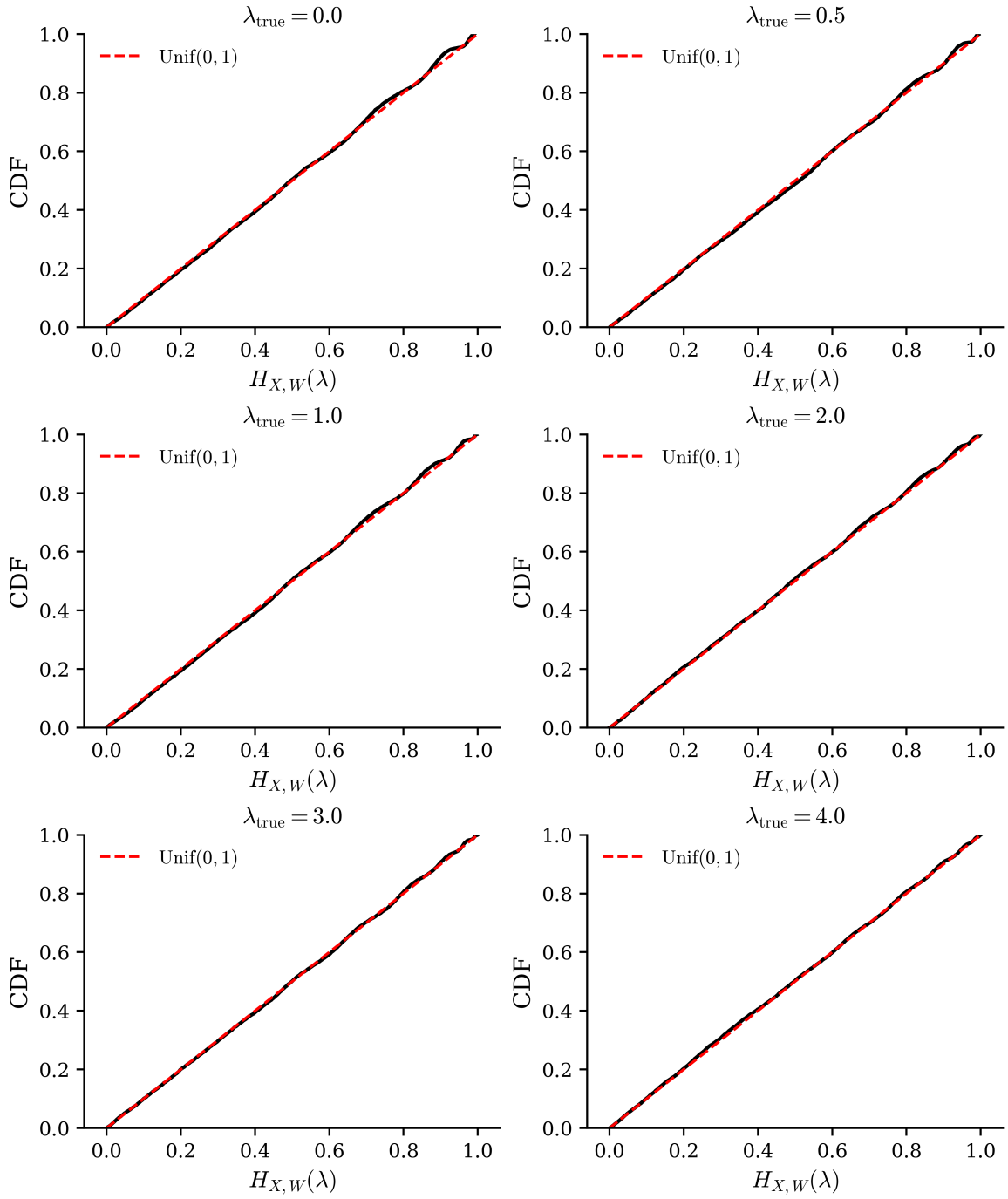

Figure 5. Empirical distribution functions of $H_{X,W}(\lambda)$ (black) compared with those of Unif(0, 1) (red) based on the random samples $X \sim \text{Poisson}(\theta)$ and $W \sim \text{Poisson}(m\varepsilon)$, where $\theta = \varepsilon + \lambda$, $\varepsilon = 3.0$, $\lambda = 0.0,\ 0.5,\ 1.0,\ 2.0,\ 3.0,\ 4.0$, and $m = 20$.

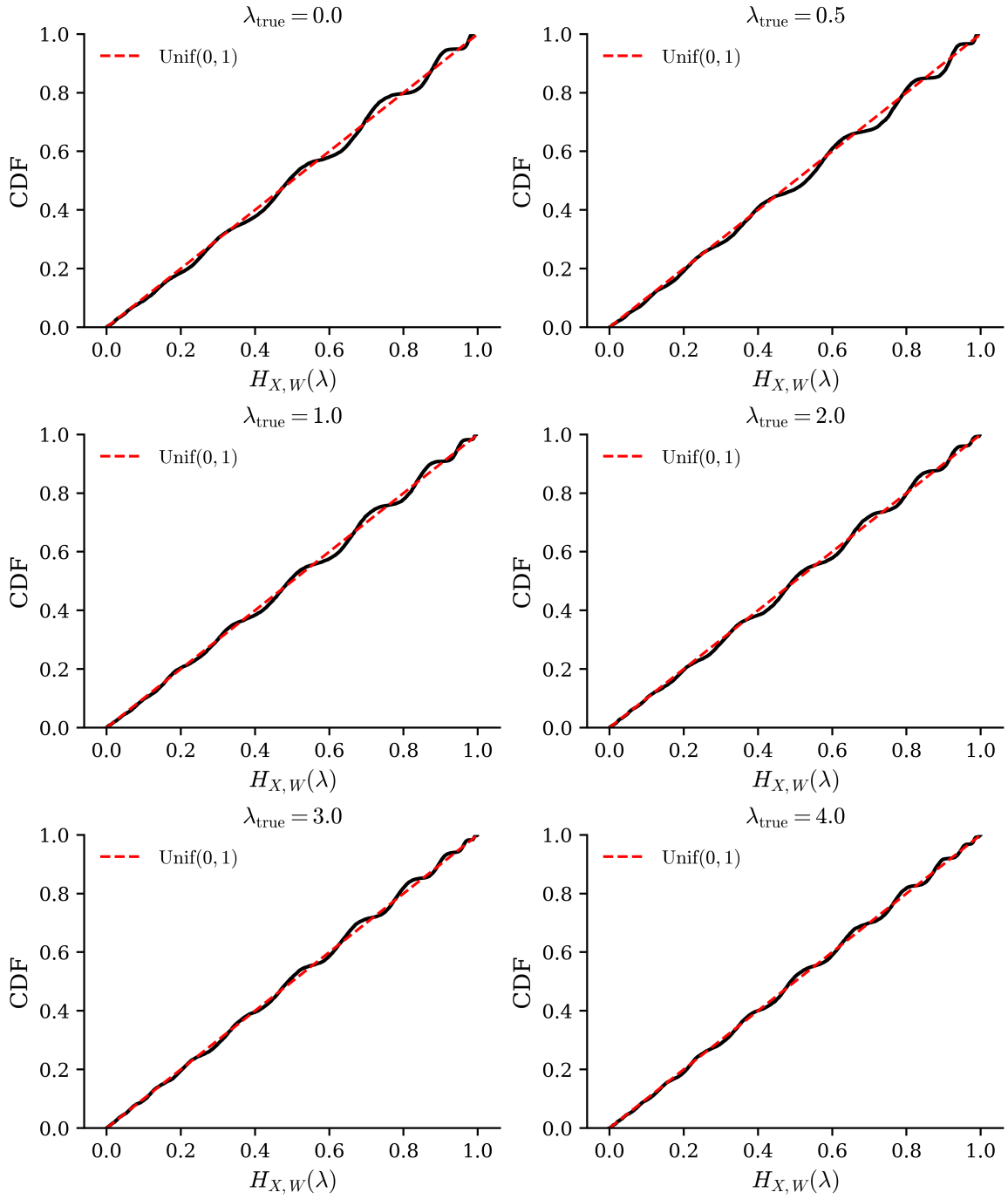

Figure 6. Empirical distribution functions of $H_{X,W}(\lambda)$ (black) compared with those of Unif(0, 1) (red) based on the random samples $X \sim \text{Poisson}(\theta)$ and $W \sim \text{Poisson}(m\varepsilon)$, where $\theta = \varepsilon + \lambda$, $\varepsilon = 3.0$, $\lambda = 0.0,\ 0.5,\ 1.0,\ 2.0,\ 3.0,\ 4.0$, and $m = 50$.

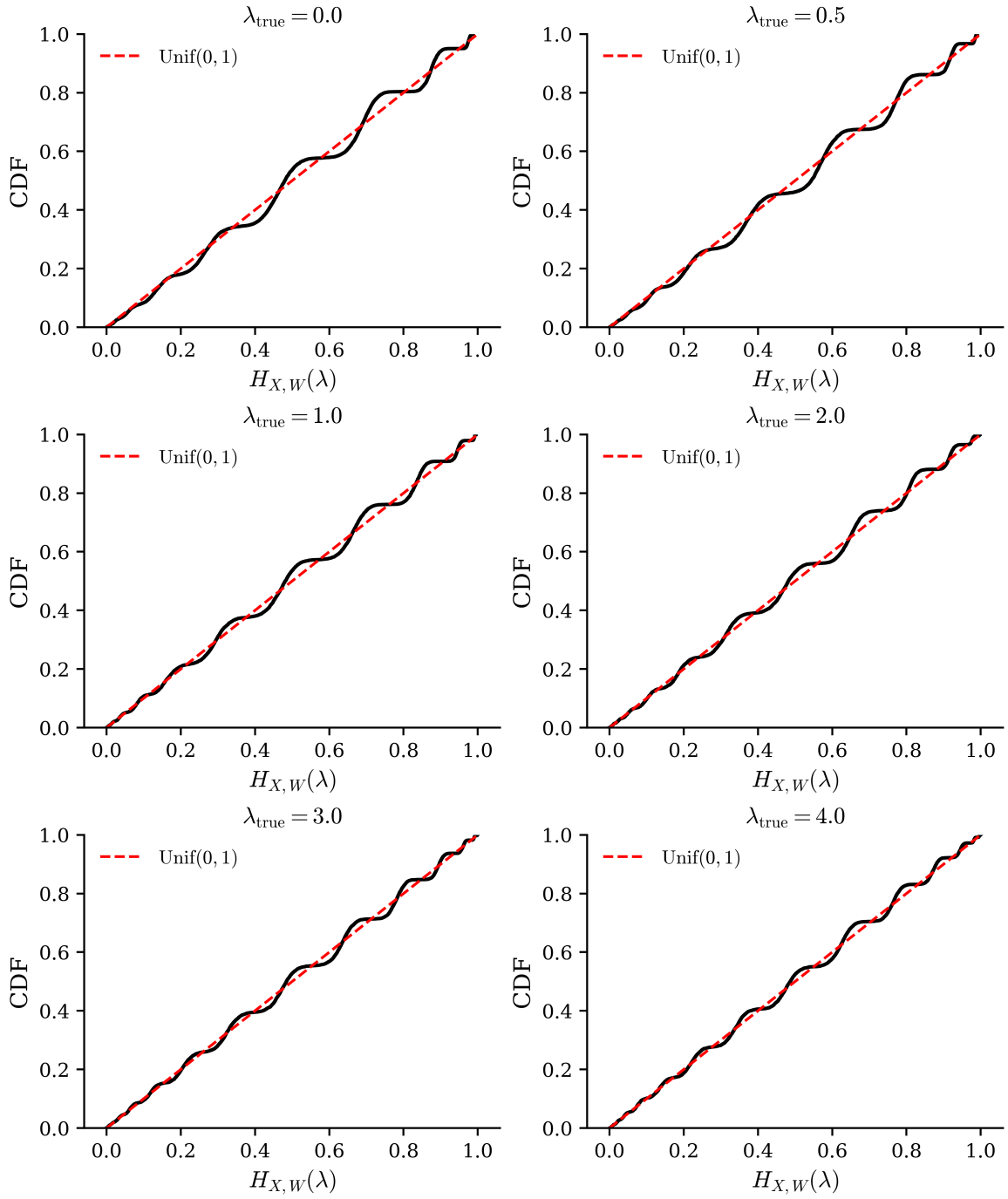

Figure 7. Empirical distribution functions of $H_{X,W}(\lambda)$ (black) compared with those of Unif(0, 1) (red) based on the random samples $X \sim \text{Poisson}(\theta)$ and $W \sim \text{Poisson}(m\varepsilon)$, where $\theta = \varepsilon + \lambda$, $\varepsilon = 3.0$, $\lambda = 0.0,\ 0.5,\ 1.0,\ 2.0,\ 3.0,\ 4.0$, and $m = 100$.

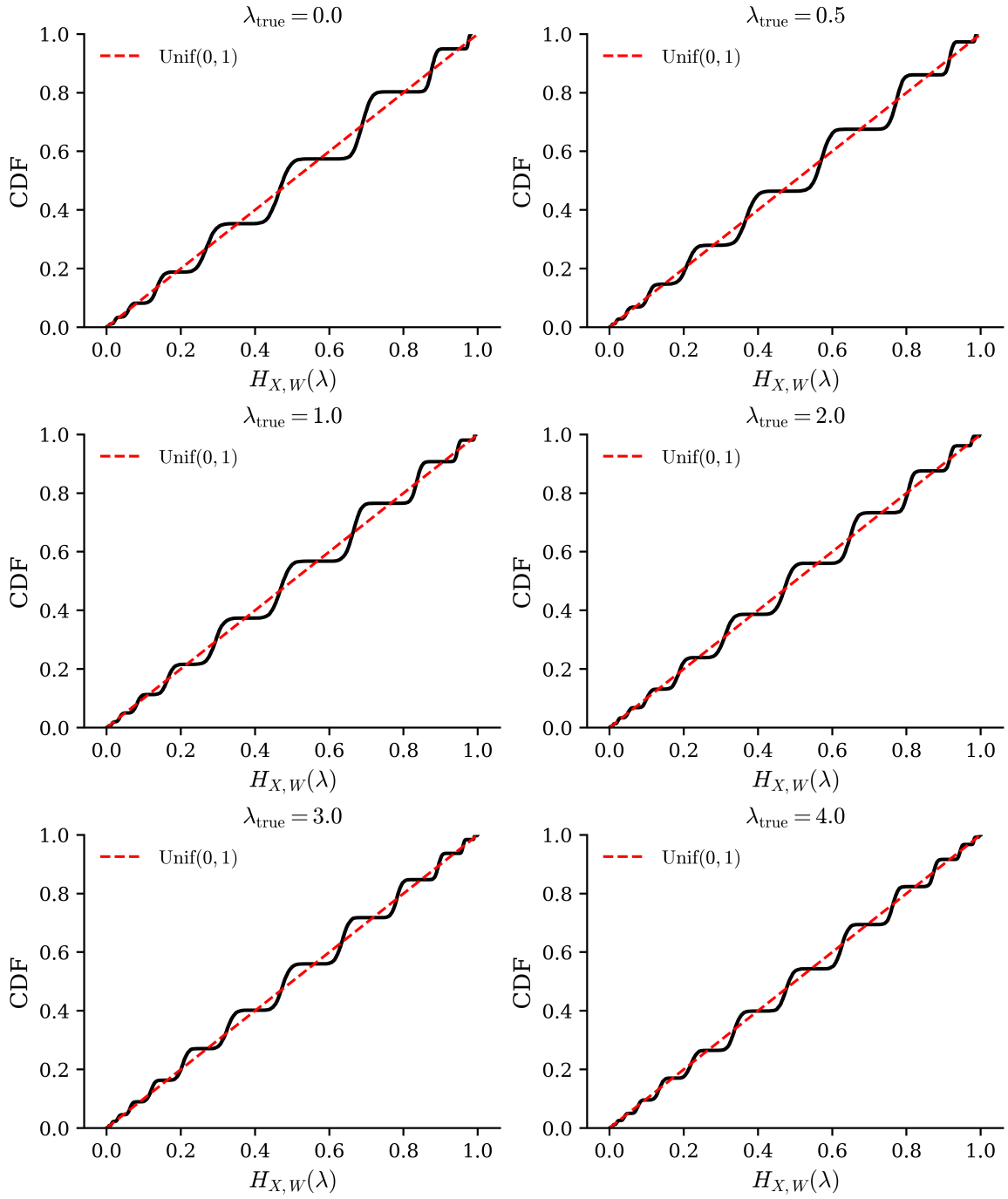

Figure 8. Empirical distribution functions of $H_{X,W}(\lambda)$ (black) compared with those of Unif(0, 1) (red) based on the random samples $X \sim \text{Poisson}(\theta)$ and $W \sim \text{Poisson}(m\varepsilon)$, where $\theta = \varepsilon + \lambda$, $\varepsilon = 3.0$, $\lambda = 0.0,\ 0.5,\ 1.0,\ 2.0,\ 3.0,\ 4.0$, and $m = 300$.

5. Real data analysis

In this section, we illustrate the application of the proposed intervals with two real examples.

Example 5.1 (Neutrino mass inference)**.** Neutrino mass is one of the core unresolved issues in particle physics. For many years, researchers have universally believed that neutrinos are massless. However, a Super-Kamiokande experiment in Japan (Fukuda et al. [16]) first confirmed that neutrinos have nonzero masses through neutrino oscillation. In recent years, KATRIN experiments in Germany (Nucciotti [17]) have improved the precision of the neutrino mass to 0.8 electron volts (eV) and further reduced it to 0.45 eV (Aker et al. [18]). Using the constrained normal mean model, we set $X = 0.45$, which serves as the observed measurement of the neutrino mass. The random variable W in our model mimics an independent auxiliary measurement of the experimental uncertainty (e.g., the variance of the background noise). Its distribution follows $W \sim \sigma^2\chi_r^2/r$, where σ^2 is the unknown variance and r is the degree of freedom. A larger r indicates a more precise uncertainty estimation. Following Zhang and Woodroofe [8], we consider $W = 0.01, 0.1, 0.5, 1.0, 5.0, 10$ to cover small, medium, and large auxiliary variabilities and $r = 5, 10, 20, 50$ to represent auxiliary information ranging from low precision to high precision. We employ the Bayesian and IM methods to compute the 90% and 95% CIs for the neutrino mass, with the corresponding results summarized in Table 3.

The 90% and 95% CIs for the neutrino mass under different W and r values are listed in Table 3. The lengths of both the Bayesian and the IM intervals decrease as r increases. In most cases, the IM interval is shorter than the Bayesian interval. However, in certain scenarios, (e.g., $W = 10$ and $r = 5$), the Bayesian interval is considerably shorter (e.g., 90% CI $[0.0000, 3.0651]$ versus the IM interval $[0.0000, 6.2209]$). The main reason for this phenomenon is that the Bayesian interval suffers from undercoverage, whereas the IM interval guarantees exact nominal coverage, as confirmed by Figures 1 and 2. Furthermore, since W reflects information about the nuisance parameter σ, when W is large and X is small (e.g., $W = 10$ and $X = 0.45$), the corresponding value of σ is also large, which increases the influence of σ on the interval estimation of θ and makes it more difficult to obtain an accurate estimate of θ; consequently, a longer interval width is required to correctly capture the true value of θ. Finally, the IM plausibility function provides interpretable evidence for each parameter value θ. For example, as shown in Figure 9, when $X = 0.45$, $W = 1.0$, $r = 10$, and $\alpha = 0.1$, every θ in the IM interval $[0.2324, 0.6659]$ has a plausibility greater than 0.1, meaning that none of these values can be rejected at the 0.10 significance level. This feature is unique to the IM framework and is not available from Bayesian or frequentist intervals. Therefore, we recommend the use of the IM CI for neutrino mass detection.

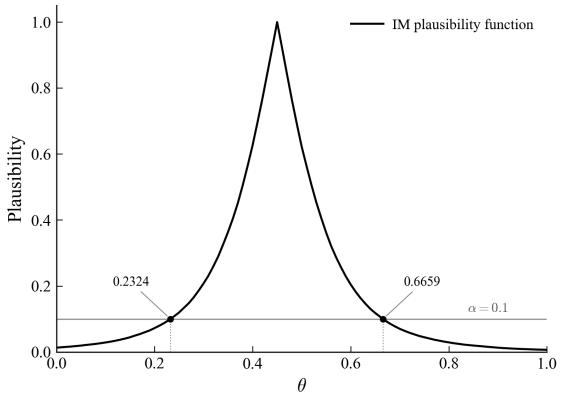

Figure 9. Plausibility function for the constrained normal mean when $X = 0.45$, $W = 1.0$, and $r = 10$.

Table 3. Levels 0.90 and 0.95 Bayesian and IM CIs and their interval widths under different combinations of observations (X, W, r).

Parameter		Level 0.90 CIs and interval widths		Level 0.95 CIs and interval widths	
W	r	Bayesian	IM	Bayesian	IM
0.01	5	[0.3599, 0.5401]	[0.4442, 0.4558]	[0.3351, 0.5649]	[0.4415, 0.4585]
		0.1802	0.0116	0.2298	0.0171
	10	[0.3927, 0.5073]	[0.4478, 0.4522]	[0.3795, 0.5205]	[0.4471, 0.4529]
		0.1146	0.0043	0.1409	0.0058
	20	[0.4114, 0.4886]	[0.4491, 0.4509]	[0.4034, 0.4966]	[0.4488, 0.4512]
		0.0771	0.0019	0.0933	0.0024
	50	[0.4263, 0.4737]	[0.4497, 0.4503]	[0.4216, 0.4784]	[0.4496, 0.4504]
		0.0474	0.0007	0.0568	0.0008
0.10	5	[0.1766, 0.7234]	[0.3919, 0.5078]	[0.1106, 0.7894]	[0.3643, 0.5354]
		0.5468	0.1159	0.6788	0.1711
	10	[0.2691, 0.6309]	[0.4283, 0.4718]	[0.2278, 0.6722]	[0.4211, 0.4789]
		0.3619	0.0434	0.4444	0.0578
	20	[0.3280, 0.5720]	[0.4406, 0.4595]	[0.3025, 0.5975]	[0.4381, 0.4619]
		0.2439	0.0189	0.2950	0.0238
	50	[0.3751, 0.5249]	[0.4465, 0.4535]	[0.3602, 0.5398]	[0.4458, 0.4542]
		0.1499	0.0070	0.1797	0.0085
0.50	5	[0.0000, 0.9446]	[0.1590, 0.7385]	[0.0000, 1.1153]	[0.0235, 0.8769]
		0.9446	0.5795	1.1153	0.8533
	10	[0.0836, 0.8164]	[0.3416, 0.5588]	[0.0212, 0.8788]	[0.3059, 0.5939]
		0.7328	0.2172	0.8576	0.2880
	20	[0.1811, 0.7189]	[0.4028, 0.4971]	[0.1274, 0.7726]	[0.3907, 0.5094]
		0.5378	0.0943	0.6452	0.1187
	50	[0.2824, 0.6176]	[0.4326, 0.4674]	[0.2492, 0.6508]	[0.4288, 0.4712]
		0.3352	0.0348	0.4017	0.0424
1.00	5	[0.0000, 1.1791]	[0.0000, 1.0286]	[0.0000, 1.4212]	[0.0000, 1.3117]
		1.1791	1.0286	1.4212	1.3117
	10	[0.0000, 0.9042]	[0.2324, 0.6659]	[0.0000, 1.0419]	[0.1624, 0.7374]
		0.9042	0.4335	1.0419	0.5750
	20	[0.0930, 0.8070]	[0.3556, 0.5441]	[0.0338, 0.8662]	[0.3313, 0.5689]
		0.7140	0.1886	0.8324	0.2376
	50	[0.2138, 0.6862]	[0.4153, 0.4849]	[0.1675, 0.7325]	[0.4076, 0.4923]
		0.4724	0.0696	0.5651	0.0847
5.00	5	[0.0000, 2.2436]	[0.0000, 3.3612]	[0.0000, 2.7895]	[0.0000, 4.7088]
		2.2436	3.3612	2.7895	4.7088
	10	[0.0000, 1.5644]	[0.0000, 1.5318]	[0.0000, 1.8661]	[0.0000, 1.8898]
		1.5644	1.5318	1.8661	1.8898
	20	[0.0070, 1.1758]	[0.0000, 0.9216]	[0.0000, 1.3687]	[0.0000, 1.0454]
		1.1758	0.9216	1.3687	1.0454
	50	[0.0123, 0.8877]	[0.2763, 0.6234]	[0.0000, 0.9933]	[0.2386, 0.6620]
		0.8755	0.3472	0.9933	0.4234
10.00	5	[0.0000, 3.0651]	[0.0000, 6.2209]	[0.0000, 3.8400]	[0.0000, 8.9531]
		3.0651	6.2209	3.8400	8.9531
	10	[0.0000, 2.0818]	[0.0000, 2.6215]	[0.0000, 2.5055]	[0.0000, 3.3284]
		2.0818	2.6215	2.5055	3.3284
	20	[0.0000, 1.5173]	[0.0000, 1.3945]	[0.0000, 1.7861]	[0.0000, 1.6341]
		1.5173	1.3945	1.7861	1.6341
	50	[0.0000, 1.0755]	[0.1020, 0.7975]	[0.0000, 1.2384]	[0.0268, 0.8725]
		1.0755	0.6955	1.2384	0.8457

Note: Cases in which the interval length is the shortest appear underlined.

Example 5.2 (Estimation of neutrino signal strength)**.** Approximately 100 trillion neutrinos pass through the human body each second in nature (Dawn [19]). However, these ghostly subatomic particles cannot be detected by our senses, as they interact only via weak forces, which makes their detection exceedingly challenging. To date, detecting neutrino signals remains a significant research

topic in high-energy physics. The KARMEN experiment searched for oscillations $\overline{\nu}_\mu \to \overline{\nu}_e$ using a liquid scintillator detector. The expected background rate $\varepsilon = 3$ was estimated from cosmic-ray neutron fluxes and the detector's response, as described by Eitel and Zeitnitz [20]. The observed total count $X = B + S$ includes both the background B and the possible signal S. In low-probability cases, X can be 0 or 1, which means that the observed events are at or below the expected background. In such situations, classical Neyman CIs may be empty, making it particularly important to construct accurate CIs for low-probability observations. Our model includes an auxiliary Poisson count $W \sim \text{Poisson}(m\varepsilon)$, which represents an independent background-monitoring measurement with a known scaling factor $m > 0$ (e.g., the ratio of exposure times between the background monitor and the main detector). A larger m indicates more precise knowledge of the background rate ε. In accordance with Zhang and Woodroofe [8], the values $m = 20, 50, 100, 300$ are chosen to cover a wide range of relative precisions, from moderate to very high. The observed W values $(10, 20, 30, 40)$ are typical realizations around the mean $m\varepsilon$.

Using the constrained Poisson mean model, we compute the Bayesian, IM, and NIM CIs for the neutrino signal strength at the 90% and 95% confidence levels, respectively, with the corresponding results summarized in Table 4. When $X = 0$, the Bayesian CI remains unchanged across different parameter settings, indicating its insensitivity to observed data. In contrast, both the IM and NIM intervals provide flexible CIs for the parameter of interest under different combinations of observed values, with the NIM interval consistently exhibiting the shortest interval length. Furthermore, when $X = 1$, the Bayesian interval has the smallest width in most cases compared with the IM and NIM intervals. Although the IM interval can be narrower than that of the Bayesian method in certain scenarios, the width of the IM interval is generally significantly greater than that of the Bayesian interval. For example, when $X = 1$, $m = 20$, and $W = 40$, the 90% Bayesian and IM CIs are $[0.00, 3.02]$ and $[0.00, 2.78]$, respectively. This is an inherent consequence of the conservative inequality-constrained model adopted by the IM method to guarantee theoretical coverage, whereas the Bayesian interval may suffer from undercoverage, a conclusion also supported by Figures 3 to 4.

As an improved version of the conservative IM interval, the NIM method consistently yields shorter widths than the IM interval for $X = 1$. Its width is closer to that of the Bayesian interval and can even be narrower in some cases. For instance, when $X = 1$, $m = 20$, and $W = 40$, the 90% Bayesian and NIM CIs are $[0.00, 2.78]$ and $[0.00, 2.14]$, respectively. Combined with the simulation analysis in Section 4.2, the NIM interval not only addresses the issues of unstable coverage performance and insensitivity to extreme data in the Bayesian interval but also mitigates the length redundancy observed in the IM interval, thereby enhancing its practical utility for estimating neutrino signal strength in low-probability observation scenarios.

Table 4. Levels 0.90 and 0.95 Bayesian, IM, and NIM CIs and their interval widths under different combinations of observations (X, W, m).

Parameter			Level 0.90 CIs and interval widths			Level 0.95 CIs and interval widths		
X	m	W	Bayesian	IM	NIM	Bayesian	IM	NIM
0	20	10	[0.00, 2.30]	[0.00, 2.50]	[0.00, 1.78]	[0.00, 3.00]	[0.00, 3.19]	[0.00, 2.49]
			2.30	2.50	1.78	3.00	3.19	2.49
		20	[0.00, 2.30]	[0.00, 2.02]	[0.00, 1.29]	[0.00, 3.00]	[0.00, 2.72]	[0.00, 1.96]
			2.30	2.02	1.29	3.00	2.72	1.96
		30	[0.00, 2.30]	[0.00, 1.53]	[0.00, 0.80]	[0.00, 3.00]	[0.00, 2.23]	[0.00, 1.47]
			2.30	1.53	0.80	3.00	2.23	1.47

Continued on the next page

Parameter			Level 0.90 CIs and interval widths			Level 0.95 CIs and interval widths		
X	m	W	Bayesian	IM	NIM	Bayesian	IM	NIM
		40	[0.00, 2.30]	[0.00, 1.05]	[0.00, 0.31]	[0.00, 3.00]	[0.00, 1.73]	[0.00, 0.98]
			2.30	1.05	0.31	3.00	1.73	0.98
0	50	10	[0.00, 2.30]	[0.00, 2.80]	[0.00, 2.08]	[0.00, 3.00]	[0.00, 3.48]	[0.00, 2.73]
			2.30	2.80	2.08	3.00	3.48	2.73
		20	[0.00, 2.30]	[0.00, 2.60]	[0.00, 1.89]	[0.00, 3.00]	[0.00, 3.30]	[0.00, 2.53]
			2.30	2.60	1.89	3.00	3.30	2.53
		30	[0.00, 2.30]	[0.00, 2.41]	[0.00, 1.69]	[0.00, 3.00]	[0.00, 3.10]	[0.00, 2.34]
			2.30	2.41	1.69	3.00	3.10	2.34
		40	[0.00, 2.30]	[0.00, 2.21]	[0.00, 1.49]	[0.00, 3.00]	[0.00, 2.90]	[0.00, 2.14]
			2.30	2.21	1.49	3.00	2.90	2.14
0	100	10	[0.00, 2.30]	[0.00, 2.89]	[0.00, 2.20]	[0.00, 3.00]	[0.00, 3.58]	[0.00, 2.83]
			2.30	2.89	2.20	3.00	3.58	2.83
		20	[0.00, 2.30]	[0.00, 2.79]	[0.00, 2.10]	[0.00, 3.00]	[0.00, 3.50]	[0.00, 2.73]
			2.30	2.79	2.10	3.00	3.50	2.73
		30	[0.00, 2.30]	[0.00, 2.70]	[0.00, 2.00]	[0.00, 3.00]	[0.00, 3.39]	[0.00, 2.63]
			2.30	2.70	2.00	3.00	3.39	2.63
		40	[0.00, 2.30]	[0.00, 2.59]	[0.00, 1.89]	[0.00, 3.00]	[0.00, 3.30]	[0.00, 2.53]
			2.30	2.59	1.89	3.00	3.30	2.53
0	300	10	[0.00, 2.30]	[0.00, 2.96]	[0.00, 2.26]	[0.00, 3.00]	[0.00, 3.66]	[0.00, 2.90]
			2.30	2.96	2.26	3.00	3.66	2.90
		20	[0.00, 2.30]	[0.00, 2.93]	[0.00, 2.23]	[0.00, 3.00]	[0.00, 3.64]	[0.00, 2.87]
			2.30	2.93	2.23	3.00	3.64	2.87
		30	[0.00, 2.30]	[0.00, 2.90]	[0.00, 2.20]	[0.00, 3.00]	[0.00, 3.58]	[0.00, 2.83]
			2.30	2.90	2.20	3.00	3.58	2.83
		40	[0.00, 2.30]	[0.00, 2.86]	[0.00, 2.17]	[0.00, 3.00]	[0.00, 3.55]	[0.00, 2.80]
			2.30	2.86	2.17	3.00	3.55	2.80
1	20	10	[0.00, 3.51]	[0.00, 4.25]	[0.00, 3.62]	[0.00, 4.36]	[0.00, 5.07]	[0.00, 4.38]
			3.51	4.25	3.62	4.36	5.07	4.38
		20	[0.00, 3.29]	[0.00, 3.77]	[0.00, 3.13]	[0.00, 4.13]	[0.00, 4.59]	[0.00, 3.90]
			3.29	3.77	3.13	4.13	4.59	3.90
		30	[0.00, 3.14]	[0.00, 3.28]	[0.00, 2.63]	[0.00, 3.97]	[0.00, 4.12]	[0.00, 3.42]
			3.14	3.28	2.63	3.97	4.12	3.42
		40	[0.00, 3.02]	[0.00, 2.78]	[0.00, 2.14]	[0.00, 3.85]	[0.00, 3.61]	[0.00, 2.93]
			3.02	2.78	2.14	3.85	3.61	2.93
1	50	10	[0.00, 3.70]	[0.00, 4.55]	[0.00, 3.92]	[0.00, 4.56]	[0.00, 5.37]	[0.00, 4.68]
			3.70	4.55	3.92	4.56	5.37	4.68
		20	[0.00, 3.57]	[0.00, 4.35]	[0.00, 3.72]	[0.00, 4.42]	[0.00, 5.17]	[0.00, 4.48]
			3.57	4.35	3.72	4.42	5.17	4.48
		30	[0.00, 3.45]	[0.00, 4.15]	[0.00, 3.53]	[0.00, 4.30]	[0.00, 4.98]	[0.00, 4.29]
			3.45	4.15	3.53	4.30	4.98	4.29
		40	[0.00, 3.36]	[0.00, 3.95]	[0.00, 3.33]	[0.00, 4.21]	[0.00, 4.79]	[0.00, 4.09]
			3.36	3.95	3.33	4.21	4.79	4.09
1	100	10	[0.00, 3.79]	[0.00, 4.64]	[0.00, 4.02]	[0.00, 4.64]	[0.00, 5.48]	[0.00, 4.79]
			3.79	4.64	4.02	4.64	5.48	4.79
		20	[0.00, 3.71]	[0.00, 4.55]	[0.00, 3.92]	[0.00, 4.56]	[0.00, 5.38]	[0.00, 4.69]
			3.71	4.55	3.92	4.56	5.38	4.69
		30	[0.00, 3.63]	[0.00, 4.45]	[0.00, 3.82]	[0.00, 4.49]	[0.00, 5.28]	[0.00, 4.59]
			3.63	4.45	3.82	4.49	5.28	4.59
		40	[0.00, 3.57]	[0.00, 4.35]	[0.00, 3.72]	[0.00, 4.42]	[0.00, 5.16]	[0.00, 4.49]
			3.57	4.35	3.72	4.42	5.16	4.49
1	300	10	[0.05, 3.88]	[0.01, 4.71]	[0.07, 4.08]	[0.01, 4.73]	[0.00, 5.54]	[0.17, 4.86]
			3.83	4.71	4.01	4.72	5.54	4.84
		20	[0.02, 3.84]	[0.00, 4.69]	[0.03, 4.05]	[0.00, 4.68]	[0.00, 5.50]	[0.00, 4.83]
			3.82	4.69	4.02	4.68	5.50	4.83
		30	[0.00, 3.79]	[0.00, 4.64]	[0.00, 4.02]	[0.00, 4.65]	[0.00, 5.46]	[0.00, 4.79]
			3.79	4.64	4.02	4.65	5.46	4.79
		40	[0.00, 3.76]	[0.00, 4.61]	[0.00, 3.99]	[0.00, 4.62]	[0.00, 5.44]	[0.00, 4.76]
			3.76	4.61	3.99	4.62	5.44	4.76

Note: Cases in which the interval length is the shortest appear underlined.

To better understand the difference between the IM and NIM CIs, Figure 10 shows their plausibility functions. By setting a significance level $\alpha \in (0, 1)$, the corresponding CIs of the two methods can be straightforwardly obtained from the plausibility function. For example, when $\alpha = 0.05$, $X = 0$, $W = 10$, and $m = 20$, the IM and NIM CIs are $[0.00, 3.19]$ and $[0.00, 2.49]$, respectively. Clearly, the NIM interval is shorter than the IM interval. More importantly, these CIs derived under the inferential model framework have a desirable property: every point inside the interval has an intuitive plausibility measure. In contrast, the Bayesian CI cannot provide such a measure and is generally longer than the NIM interval. Therefore, we recommend the use of the NIM method for estimating neutrino signal strength in practical applications.

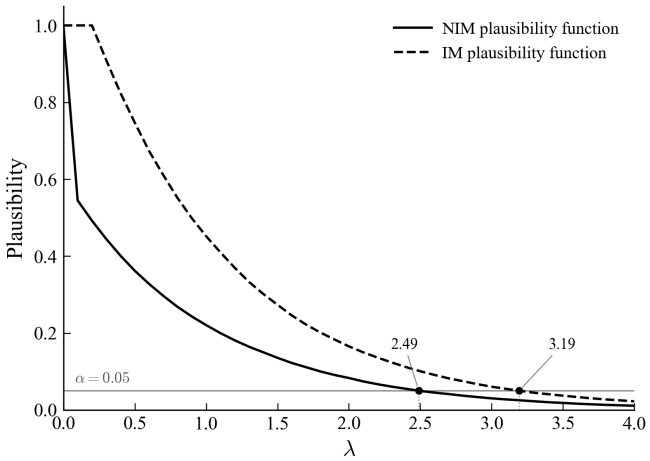

Figure 10. Plausibility function for the constrained Poisson mean when $X = 0$, $W = 10$, and $m = 20$.

6. Discussion

Building upon the IM framework, this study proposes interval estimation methods for constrained parameters in normal and Poisson distributions with unknown nuisance parameters. The core contribution lies in addressing key shortcomings of conventional approaches. Compared with Bayesian methods, which exhibit undercoverage in the normal and Poisson cases, the IM intervals integrate parameter constraints with PRSs to ensure stable coverage closely aligned with nominal levels. The NIM intervals further correct for the discreteness of Poisson distributions via random weighting, improving inferential efficiency while achieving superior interval lengths, particularly under low-count scenarios where they significantly outperform Bayesian intervals. Additionally, the proposed IM-based methods are prior-free, and each parameter value within the interval is accompanied by a likelihood-based measure, providing more physically interpretable inference for fields such as high-energy physics.

From a practical perspective, the IM and NIM methods demonstrate considerable value in cutting-edge scientific problems such as neutrino mass measurement and signal strength estimation. They not only resolve issues where traditional intervals may be empty or overly conservative under low-count observations, but also offer refined characterization of parameter uncertainty through plausibility functions, thereby delivering more reliable statistical support for experimental design and result

interpretation. As a prior-free probabilistic inferential framework, this approach not only enriches the methodological system for constrained parameter inference but also offers an effective new pathway for analyzing complex data in fields such as high-energy physics and astronomical observation.

In practice, the appropriate use of the IM and NIM intervals should be guided by the specific inferential goal and the characteristics of the data. For the constrained normal mean, we recommend the IM interval as a more reliable alternative to the Bayesian interval. Not only does it guarantee nominal coverage, but its interval length is also shorter than that of the Bayesian interval when the auxiliary observation W is small or the degrees of freedom r are large. Therefore, the IM interval is particularly suitable for problems such as neutrino mass detection, where the true mass is close to zero. For the constrained Poisson mean, we recommend the NIM interval. Although the IM interval has conservative coverage and is thus a more reliable interval estimation method, its interval length is often significantly longer than those of the Bayesian and NIM intervals. In contrast, the NIM interval achieves coverage closest to the nominal level, and its interval length is often shorter than those of both the Bayesian and IM intervals in extreme observation cases such as $X = 0$. Hence, the NIM interval is particularly well-suited for applications such as neutrino signal detection.

Although the proposed IM and NIM methods outperform the Bayesian method, they are not yet optimal for inference on constrained parameters. For the constrained normal mean, the Bayesian method is inaccurate. In contrast, the IM method achieves exact coverage when $\theta > 0$. However, at $\theta = 0$, its coverage probability is slightly higher than the nominal confidence level. Therefore, optimal interval estimation methods for the parameter at the boundary value are still needed. For the constrained Poisson mean, although the IM method addresses the undercoverage issue of the Bayesian method, the discreteness of the Poisson distribution also makes the IM CIs somewhat conservative. To overcome the impact of data discreteness on IM inference, the NIM method based on random weighting achieves more accurate coverage than both the Bayesian and IM methods when the value of m is small. Notably, when m is large, $H_{X,W}(\lambda)$ violates the uniformity condition required by Theorem 3.4, resulting in NIM intervals that are shorter in practice than IM intervals and thus may fail to guarantee nominal coverage. Therefore, to obtain CIs with guaranteed nominal coverage, we recommend using longer IM intervals. If one seeks a better CI between the undercovering Bayesian interval and the conservative IM interval, NIM provides a viable alternative. Future research may not only provide better solutions for incorporating parameter constraints into IM inference but also extend the proposed IM and NIM methods to other constrained statistical models and inference problems involving multiple parameter constraints.

7. Conclusions

On the basis of the IM framework, this paper develops prior-free interval estimation for constrained normal and Poisson means with unknown nuisance parameters. The proposed IM interval achieves exact nominal coverage for the constrained normal mean, whereas for the constrained Poisson mean, the IM interval is conservative but valid, and the NIM interval further improves coverage accuracy and shortens interval length via random weighting. Real data examples on neutrino mass and signal strength demonstrate the practical advantages of these methods over Bayesian alternatives. Overall, the IM and NIM intervals provide reliable, interpretable, and prior-free uncertainty quantification for constrained parameter problems, with the IM interval recommended for normal cases and the NIM

interval for Poisson cases, especially under low–count observations.

Author contributions

Hezhi Lu: Conceptualization, investigation, supervision, methodology, validation, project administration, writing–review and editing, funding acquisition; Qijun Wu: Data curation, data analysis, software, formal analysis, visualization, writing–original draft. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work was supported by the Guangdong Basic and Applied Basic Research Foundation (2023A1515110661).

Conflict of interest

The authors declare no conflict of interest.

References

1. G. J. Feldman, R. D. Cousins, Unified approach to the classical statistical analysis of small signals, *Phys. Rev. D*, **57** (1998), 3873–3889. https://doi.org/10.1103/PhysRevD.57.3873
2. M. Mandelkern, J. Schultz, The statistical analysis of Gaussian and Poisson signals near physical boundaries, *J. Math. Phys.*, **41** (2000), 5701–5709. https://doi.org/10.1063/1.533433
3. C. Giunti, New ordering principle for the classical statistical analysis of Poisson processes with background, *Phys. Rev. D*, **59** (1999), 053001. https://doi.org/10.1103/PhysRevD.59.053001
4. M. Mandelkern, Setting confidence intervals for bounded parameters, *Statist. Sci.*, **17** (2002), 149–172. https://doi.org/10.1214/ss/1030550859
5. D. A. S. Fraser, N. Reid, A. C. M. Wong, Inference for bounded parameters, *Phys. Rev. D*, **69** (2004), 033002. https://doi.org/10.1103/PhysRevD.69.033002
6. Y. Zhu, Upper limit for Poisson variable incorporating systematic uncertainties by Bayesian approach, *Nucl. Instrum. Methods Phys. Res. A*, **578** (2007), 322–328. https://doi.org/10.1016/j.nima.2007.05.116
7. B. P. Roe, M. B. Woodroofe, Setting confidence belts, *Phys. Rev. D*, **63** (2000), 013009. https://doi.org/10.1103/PhysRevD.63.013009
8. T. Zhang, M. Woodroofe, Credible and confidence sets for restricted parameter spaces, *J. Stat. Plan. Infer.*, **115** (2003), 479–490. https://doi.org/10.1016/S0378-3758(02)00170-2

9. D. E. Leaf, C. Liu, Inference about constrained parameters using the elastic belief method, *Int. J. Approx. Reason.*, **53** (2012), 709–727. https://doi.org/10.1016/j.ijar.2012.02.003
10. H. Lu, H. Jin, Z. Wang, Y. Li, Confidence intervals for a Poisson parameter with background, *Commun. Stat. Theory Methods*, **52** (2023), 6794–6805. https://doi.org/10.1080/03610926.2022.2033268
11. C. Chen, S. Chen, S. Wang, D. Wang, Y. Chen, Z. Zeng, Fiducial inference framework for restricted parameter spaces: Poisson mean with background, *BMC Med. Res. Methodol.*, **26** (2026), 48. https://doi.org/10.1186/s12874-026-02812-5
12. A. P. Dempster, The Dempster–Shafer calculus for statisticians, *Int. J. Approx. Reason.*, **48** (2008), 365–377. https://doi.org/10.1016/j.ijar.2007.03.004
13. G. Shafer, *A mathematical theory of evidence*, Princeton: Princeton University Press, 1976. https://doi.org/10.1515/9780691214696
14. R. Martin, C. Liu, Inferential models: A framework for prior-free posterior probabilistic inference, *J. Am. Stat. Assoc.*, **108** (2013), 301–313. https://doi.org/10.1080/01621459.2012.747960
15. H. Lu, H. Jin, Z. Wang, C. Chen, Y. Lu, Prior-free probabilistic interval estimation for binomial proportion, *TEST*, **28** (2019), 522–542. https://doi.org/10.1007/s11749-018-0588-0
16. Y. Fukuda, T. Hayakawa, E. Ichihara, K. Inoue, K. Ishihara, H. Ishino, et al., Evidence for oscillation of atmospheric neutrinos, *Phys. Rev. Lett.*, **81** (1998), 1562–1567. https://doi.org/10.1103/PhysRevLett.81.1562
17. A. Nucciotti, Still too small to be measured, *Nat. Phys.*, **18** (2022), 128–129. https://doi.org/10.1038/s41567-021-01495-7
18. M. Aker, D. Batzler, A. Beglarian, J. Behrens, J. Beisenkötter, M. Biassoni, et al., Direct neutrino-mass measurement based on 259 days of KATRIN data, *Science*, **388** (2025), 180–185. https://doi.org/10.1126/science.adq9592
19. D. Levy, Neutrino experiments explore the unknown with ORNL expertise, equipment, In: *Oak Ridge national laboratory review*, **50** (2017), 20–21. Available from: `https://www.ornl.gov/content/ornl-review-v50n1`
20. K. Eitel, B. Zeitnitz, KARMEN collaboration, The search for neutrino oscillations $\overline{\nu_\mu} \to \overline{\nu_e}$ with KARMEN, *Nucl. Phys. B Proc. Suppl.*, **77** (1999), 212–219. https://doi.org/10.1016/S0920-5632(99)00420-X

Appendix A: Technical details

Proof of Lemma 3.1. For fixed $X, W \in \{0, 1, \ldots\}$, $m > 0$, and $b = 1$, let U and $\widetilde{U}$ be independent Unif(0, 1) variables. Since $G_{a,1}^{-1}(\cdot)$ is monotonically increasing in a, we have

$$G_{X,1}^{-1}(1-U) \le G_{X+1,1}^{-1}(1-U), \qquad G_{W,1}^{-1}(1-\widetilde{U}) \le G_{W+1,1}^{-1}(1-\widetilde{U}).$$

Hence, $L_{X,W}(U, \widetilde{U}) \le R_{X,W}(U, \widetilde{U})$ holds almost everywhere. Consequently, for any $t \in \mathbb{R}$,

$$\left\{(U, \widetilde{U}) : R_{X,W}(U, \widetilde{U}) \le t\right\} \subseteq \left\{(U, \widetilde{U}) : L_{X,W}(U, \widetilde{U}) \le t\right\}.$$

Hence, $K_{X+1,W}(t) \leq K_{X,W+1}(t)$. Moreover, $K^{-1}_{X,W+1}(v) \leq K^{-1}_{X+1,W}(v)$ for $0 < v < 1$.

Next, we demonstrate that (3.9) is equivalent to (3.10). On one hand, since $G^{-1}_{X+1,1}(1-U)$ and $G^{-1}_{W,1}(1-\widetilde{U})$ are independent continuous random variables, $R_{X,W}(U,\widetilde{U})$ has a continuous distribution. Similarly, the distribution function of $L_{X,W}(U,\widetilde{U})$ is also continuous. Therefore, for any $\lambda \geq 0$, we have

$$P_{U,\widetilde{U}}\left\{R_{X,W}(U,\widetilde{U}) < \lambda\right\} = P_{U,\widetilde{U}}\left\{R_{X,W}(U,\widetilde{U}) \leq \lambda\right\} = K_{X+1,W}(\lambda)$$

and

$$\begin{aligned} &P_{U,\widetilde{U}}\left\{L_{X,W}(U,\widetilde{U}) < \lambda < R_{X,W}(U,\widetilde{U})\right\} \\ &= P_{U,\widetilde{U}}\left\{L_{X,W}(U,\widetilde{U}) < \lambda\right\} - P_{U,\widetilde{U}}\left\{R_{X,W}(U,\widetilde{U}) < \lambda\right\} \\ &= K_{X,W+1}(\lambda) - K_{X+1,W}(\lambda). \end{aligned}$$

On the other hand, for any $\lambda \geq 0$ and $V \sim \text{Unif}(0,1)$, by the quantile property of continuous distribution functions,

$$\begin{aligned} &P_V\left\{K^{-1}_{X,W+1}(V) < \lambda < K^{-1}_{X+1,W}(V)\right\} \\ &= P_V\left\{K_{X+1,W}(\lambda) < V < K_{X,W+1}(\lambda)\right\} \\ &= K_{X,W+1}(\lambda) - K_{X+1,W}(\lambda). \end{aligned}$$

Thus,

$$P_{U,\widetilde{U}}\left\{L_{X,W}(U,\widetilde{U}) < \lambda < R_{X,W}(U,\widetilde{U})\right\} = P_V\left\{K^{-1}_{X,W+1}(V) < \lambda < K^{-1}_{X+1,W}(V)\right\}.$$

Hence, the proof is complete. □

Proof of Theorem 3.1. To prove Theorem 3.1, it suffices to show that the IM is valid, which has already been established in Theorem 2 of Martin and Liu [14]. □

Proof of Theorem 3.2. From (3.4), $F\left(\frac{X-\theta}{\sqrt{W}}\right) \sim \text{Unif}(0,1)$. When $\theta_0 = 0$, we have

$$\begin{aligned} &\sup_{\theta\in A} P_{(X,W)|\theta,\sigma^2}\left\{pl_{X,W}(A) \leq \alpha \mid \theta_0 = 0\right\} \\ &= \sup_{\theta\in A} P_{(X,W)|\theta,\sigma^2}\left\{F\left(\frac{X}{\sqrt{W}}\right) \geq 1-\frac{\alpha}{2}, \frac{1}{2} < F\left(\frac{X}{\sqrt{W}}\right) \leq 1 \mid \theta_0 = 0\right\} \\ &= \frac{\alpha}{2}. \end{aligned}$$

When $\theta_0 > 0$, we have

$$\begin{aligned} &\sup_{\theta\in A} P_{(X,W)|\theta,\sigma^2}\left\{pl_{X,W}(A) \leq \alpha \mid \theta = \theta_0\right\} \\ &= \sup_{\theta\in A}\left[P_{(X,W)|\theta,\sigma^2}\left\{F\left(\frac{X-\theta_0}{\sqrt{W}}\right) \leq \frac{\alpha}{2}, F\left(\frac{X-\theta_0}{\sqrt{W}}\right) < \frac{1}{2} \mid \theta = \theta_0\right\}\right. \\ &\quad \left. + P_{(X,W)|\theta,\sigma^2}\left\{F\left(\frac{X-\theta_0}{\sqrt{W}}\right) \geq 1-\frac{\alpha}{2}, F\left(\frac{X-\theta_0}{\sqrt{W}}\right) \geq \frac{1}{2} \mid \theta = \theta_0\right\}\right] \\ &= \alpha. \end{aligned}$$

Therefore,

$$\sup_{\theta\in A} P_{(X,W)|\theta,\sigma^2}\left\{pl_{X,W}(A) \leq \alpha\right\} \leq \alpha.$$

By Theorem 3.1, the proof is complete. □

Proof of Theorem 3.3. From (3.10), we have $K^{-1}_{X,W+1}(V) < \lambda < K^{-1}_{X+1,W}(V)$, $V \sim \text{Unif}(0,1)$. Given any $\alpha \in (0,1)$,

$$\begin{aligned}&\sup_{\lambda\in A} P_{(X,W)|\lambda,\varepsilon}\left\{pl_{X,W}(A) \le \alpha \mid \lambda_0 = 0\right\}\\ =&\sup_{\lambda\in A} P_{(X,W)|\lambda,\varepsilon}\left\{2K_{X,W+1}(0) \le \alpha, K_{X,W+1}(0) < \frac{1}{2} \mid \lambda_0 = 0\right\} = \frac{\alpha}{2},\end{aligned}$$

when $\lambda_0 = 0$, and

$$\begin{aligned}&P_{(X,W)|\lambda,\varepsilon}\left\{pl_{X,W}(A) \le \alpha \mid \lambda = \lambda_0\right\}\\ =& P_{(X,W)|\lambda,\varepsilon}\left\{2K_{X,W+1}(\lambda_0) \le \alpha, K_{X,W+1}(\lambda_0) < \frac{1}{2}\right\} + P_{(X,W)|\lambda,\varepsilon}\left\{2 - 2K_{X+1,W}(\lambda_0) \le \alpha, K_{X+1,W}(\lambda_0) > \frac{1}{2}\right\}\\ \le& P_{(X,W)|\lambda,\varepsilon}\left\{V \le \frac{\alpha}{2}, V < \frac{1}{2}\right\} + P_{(X,W)|\lambda,\varepsilon}\left\{V \ge 1 - \frac{\alpha}{2}, V > \frac{1}{2}\right\} = \alpha,\end{aligned}$$

when $\lambda_0 > 0$. Afterward, $\sup_{\lambda\in A} P_{(X,W)|\lambda,\varepsilon}\left\{pl_{X,W}(A) \le \alpha\right\} \le \alpha$. Hence, the proof is complete. □

Proof of Theorem 3.4. Given any $\alpha \in (0,1)$, since $H_{X,W}(\lambda) \sim \text{Unif}(0,1)$ for $(X,W) \sim P_{(X,W)|\lambda,\varepsilon}$ for all λ, then

$$\begin{aligned}pl_{X,W}(A) &= P_S\left\{\Theta_{X,W}(S) \cap A \ne \varnothing\right\}\\ &= P_S\left\{H_{X,W}(\lambda) \in S\right\}\\ &= P_S\left\{v \in S\right\}.\end{aligned}$$

Moreover, the PRS $S \sim P_S$ is valid; that is, $P_S\left\{v \in S\right\} \ge_{st} \text{Unif}(0,1)$.

$$\begin{aligned}\sup_{\lambda\in A} P_{(X,W)|\lambda,\varepsilon}\left\{pl_{X,W}(A) \le \alpha\right\} &= \sup_{\lambda\in A} P_{(X,W)|\lambda,\varepsilon}\left\{P_S\left\{v \in S\right\} \le \alpha\right\}\\ &\le \sup_{\lambda\in A} P_{(X,W)|\lambda,\varepsilon}\left\{\text{Unif}(0,1) \le \alpha\right\}\\ &= \alpha.\end{aligned}$$

Hence, the proof is complete. □

AIMS Press