Transactions on Fuzzy Sets and Systems (TFSS)
URL: https://sanad.iau.ir/journal/tfss/
Online ISSN: 2821-0131
Vol.5, No.1, (2026), 182-207
DOI: https://doi.org/10.71602/tfss.2026.1229395

Article Type: Original Research Article

Neutrosophic Negative Binomial Hurdle Models for Overdispersed Zero-Inflated Count Data

Ibrahim Sadok

Abstract. The Negative Binomial Hurdle (NBH) model stands as a preeminent statistical framework for analyzing overdispersed count data characterized by a preponderance of zero observations. Its two-part structure (decoupling the probability of a zero outcome from the distribution of positive counts) provides a powerful and intuitive mechanism for capturing complex data-generating processes. However, the conventional NBH model, alongside its classical counterparts, operates under the stringent assumption of precisely defined parameters, a condition often violated in real-world empirical contexts. Data is frequently imbued with inherent indeterminacy, stemming from measurement inaccuracies, ambiguous classification criteria, subjective judgment, or intrinsic system volatility. This paper introduces the Neutrosophic Negative Binomial Hurdle (NNBH) model, a novel generalization that integrates the principles of neutrosophic logic into the hurdle framework. By formulating key parameters (the hurdle probability, the count rate, or both) as neutrosophic numbers, the proposed model quantifies indeterminacy through interval-based estimates, thereby offering a more honest and robust representation of underlying uncertainties. We formally derive the model's probability mass function, likelihood, and fundamental statistical properties. Parameter estimation is achieved through an adapted Neutrosophic Maximum Likelihood Estimation (NMLE) algorithm. The model's superior flexibility and performance are unequivocally demonstrated through an extensive simulation study and a compelling real-world application to healthcare utilization data, where it yields a more nuanced and informative fit than its classical and partially neutrosophic predecessors. This work provides a critical new analytical tool for researchers and practitioners in public health, economics, and industrial engineering who must derive inferences from complex, vague, and overdispersed count processes.

AMS Subject Classification 2020: 62J12; 62P10; 62E17
Keywords and Phrases: Excess Zeros, Healthcare analytics, Hurdle models, Indeterminacy, Maximum likelihood estimation, Negative binomial regression, Neutrosophic statistics, Overdispersion.

1 Introduction

The statistical modelling of count data represents a cornerstone of empirical research across a vast spectrum of disciplines, including epidemiology, economics, ecology, and industrial quality control [?, ?, ?]. A pervasive and analytically challenging feature of such data is the frequent occurrence of an excess of zero observations, a phenomenon that systematically violates the assumptions of standard Poisson and Negative Binomial distributions, leading to significant model misspecification and biased inference [?, ?, ?]. To address this limitation, the methodological arsenal of statisticians has been enriched by two prominent classes of models: zero-inflated and hurdle formulations. While zero-inflated models, such as the Zero-Inflated Poisson (ZIP), conceptualize the data as arising from a mixture of a degenerate distribution at zero and a standard count distribution, hurdle models offer a distinct, two-part paradigm [?, ?, ?]. First introduced by Mullahy [?],

∗**Corresponding Author:** Ibrahim Sadok, Email: ibrahim.sadok@univ-bechar.dz, ORCID: 0000-0002-5366-5853
Received: 25 December 2025; **Revised**: 8 January 2026; **Accepted**: 30 January 2026; **Available Online**: 19 February 2026; **Published Online**: 7 May 2026.

How to cite: Sadok I. Neutrosophic negative binomial Hurdle models for overdispersed zero-inflated count data. *Transactions on Fuzzy Sets and Systems.* 2026; 5(1): 182-207. DOI: https://doi.org/10.71602/tfss.2026.1229395

the hurdle model posits a binary process governing the transition from a zero state to a non-zero state (the "hurdle"), followed by a truncated count distribution modelling the positive outcomes. This elegant separation often provides a more interpretable framework for processes where the decision to engage in an activity is structurally different from the level of engagement thereafter.

Notwithstanding their widespread adoption and theoretical appeal, classical hurdle models, including the Negative Binomial Hurdle (NBH) [?], are predicated on the assumption of precise and determinate parameters. This assumption proves untenable in a multitude of practical scenarios where data is intrinsically fraught with indeterminacy [?, ?]. This indeterminacy can manifest from a plethora of sources: measurement errors introduce noise and imprecision; ambiguous operational definitions lead to uncertain classification of outcomes; subjective human judgement injects variability; and inherent system fluctuations create an environment of inherent volatility. Consider, for instance, a public health study investigating the "annual number of specialist medical consultations." A reported zero could indeterminately represent a genuinely healthy individual (a "true zero") or an individual with unmet healthcare needs due to financial barriers, geographical constraints, or health illiteracy (an "indeterminate zero"). Similarly, among those who do seek care, the count of visits might be imprecisely measured or driven by an underlying severity of illness that is itself vaguely defined. Classical models, constrained by their crisp parametric foundations, are inherently incapable of quantifying this vagueness, potentially yielding overconfident and misleading conclusions. The emergent field of neutrosophic statistics, pioneered by Smarandache [?], offers a profound philosophical and mathematical framework for confronting this very challenge. Neutrosophy, an extension of fuzzy and intuitionistic logic, formally incorporates three independent components (Truth (T), Indeterminacy (I), and Falsity (F)) to model the spectrum of certainty and uncertainty. Within statistics, this translates to the representation of parameters as neutrosophic numbers, defined as intervals that encapsulate both determinate and indeterminate parts [?, ?]. This paradigm shift allows for the direct quantification of ambiguity, providing a more realistic representation of complex systems. While the literature has recently witnessed the advent of neutrosophic generalizations for several standard distributions, such as the Poisson and Negative Binomial [?, ?], the two-part hurdle framework remains a conspicuously unexplored frontier within this context.

This paper seeks to bridge this critical methodological gap by introducing the Neutrosophic Negative Binomial Hurdle (NNBH) model. Our work makes several salient contributions to the literature. Firstly, we provide a formal mathematical exposition of three distinct NNBH parameterizations, each designed to embed indeterminacy at different stages of the data-generating process, thus offering tailored solutions for various sources of ambiguity. Secondly, we derive the neutrosophic likelihood function and devise a corresponding Neutrosophic Maximum Likelihood Estimation (NMLE) procedure, complete with a comprehensive algorithmic implementation. Thirdly, we undertake a rigorous simulation study to evaluate the finite-sample performance of our estimators, establishing their consistency and asymptotic properties. Finally, we demonstrate the practical supremacy and interpretive richness of the NNBH model through an empirical application to a healthcare utilization dataset, where it demonstrably outperforms classical and simpler neutrosophic models, providing deeper insights into the inherent uncertainties of patient behavior. This model thus equips analysts with a powerful and sophisticated tool for navigating the complexities of modern, imprecise data.

2 Preliminaries

2.1 The Philosophical and Mathematical Foundations of Neutrosophic Logic

Neutrosophic logic, as a generalization of classical and fuzzy logics, operates on the fundamental premise that any proposition exists in a state of spectrum defined not only by its degrees of truth and falsity but also by a distinct degree of indeterminacy [?, ?]. Formally, let T, I, F be real standard or non-standard subsets of the interval $]^{-}0, 1^{+}[$, representing the truth, indeterminacy, and falsity membership values, respectively. This triadic structure allows for a more nuanced representation of real-world information where ambiguity is not merely a lack of knowledge but an inherent characteristic.

Within the realm of probability and statistics, this philosophy materializes through the concept of neutrosophic numbers. A neutrosophic number N is defined as $N = n_l + n_u I_N$, where n_l signifies the determinate component, n_u represents the indeterminate spread, and $I_N \in [I_L, I_U]$ is the indeterminacy interval with lower and upper bounds I_L and I_U. Consequently, N encapsulates all possible values within the interval $[n_l + n_u I_L, n_l + n_u I_U]$. A neutrosophic random variable Y_N is characterized by a probability distribution whose parameters are themselves neutrosophic numbers, leading to probabilities that are not point estimates but intervals, thereby formally incorporating indeterminacy into the inferential process [**?**, **?**].

2.2 The Classical Negative Binomial Hurdle Model

The classical Negative Binomial Hurdle (NBH) model is a two-component mixture model designed to handle count data with excess zeros and overdispersion [**?**]. Let Y be a count-valued random variable. The model is specified by its probability mass function (PMF):

$$P(Y = y) = \begin{cases} \pi & \text{for } y = 0 \quad \text{(A binary process for the hurdle)} \\ (1-\pi) \cdot \dfrac{g(y)}{1-g(0)} & \text{for } y = 1, 2, 3, \ldots \quad \text{(A zero-truncated count process)} \end{cases} \tag{1}$$

where:
- $\pi \in [0, 1]$ is the hurdle probability, representing the probability of a zero observation (i.e., not crossing the hurdle).
- $g(y)$ is the PMF of a standard Negative Binomial (NB) distribution, parameterized by a mean $\mu > 0$ and a dispersion parameter $\alpha > 0$. Its form is:

$$g(y) = \frac{\Gamma(y + \alpha^{-1})}{\Gamma(y+1)\Gamma(\alpha^{-1})} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu}\right)^{\alpha^{-1}} \left(\frac{\mu}{\alpha^{-1} + \mu}\right)^{y} \tag{2}$$

- $\dfrac{g(y)}{1-g(0)}$ is the PMF of a zero-truncated Negative Binomial distribution, which governs the positive counts conditional on the hurdle having been crossed. This formulation elegantly separates the mechanism generating the zeros from the mechanism generating the positive counts, often yielding superior interpretability in applied contexts.

3 The Neutrosophic Negative Binomial Hurdle (NNBH) Model

3.1 Model Formulations

Building upon the philosophical and mathematical foundations outlined in the previous section, we now introduce the core contribution of this work: the Neutrosophic Negative Binomial Hurdle (NNBH) model. This novel class of models is designed to incorporate indeterminacy directly into the fabric of the hurdle framework, acknowledging that real-world processes are often characterized by ambiguity at one or more stages. We propose three distinct parameterizations, each targeting a specific source of indeterminacy, thereby providing a tailored and interpretable modeling toolkit for the applied researcher.

Definition 3.1. (NNBH-I: Neutrosophic Hurdle Probability) *A random variable Y_N is said to follow an NNBH-I distribution if its neutrosophic probability mass function is given by:*

$$P_N(Y_N = y) = \begin{cases} \pi_N & ; \textit{for } y = 0 \\ (1-\pi_N) \cdot \dfrac{g(y)}{1-g(0)} & ; \textit{for } y = 1, 2, 3, \ldots \end{cases} \tag{3}$$

where $\pi_N = \pi_l + \pi_u I_{N_\pi}$ is the neutrosophic hurdle probability, with $I_{N_\pi} \in [I_{L_\pi}, I_{U_\pi}]$ representing the indeterminacy interval. The parameters of the count component, namely the mean $\mu > 0$ and the dispersion parameter $\alpha > 0$, remain crisp, classical values. The model necessitates the constraint $0 \leq \pi_l + \pi_u I_{L_\pi} \leq \pi_l + \pi_u I_{U_\pi} \leq 1$ to ensure valid probabilities across the entire indeterminacy spectrum.

Interpretation and Applicability: The NNBH-I model is particularly potent in scenarios where the primary source of ambiguity lies in the binary decision process that determines whether a count is zero. The interval π_N quantitatively captures the uncertainty in classifying an observation as a structural zero. For instance, in a study on insurance claims, a zero could indeterminately represent a claim-free policyholder or a policyholder who did not file a claim for a minor incident due to a high deductible. The neutrosophic interval π_N encapsulates this spectrum of possibility, offering a more nuanced view than a single, potentially misleading, point estimate.

Definition 3.2. (NNBH-II: Neutrosophic Count Process) *A random variable Y_N follows an NNBH-II distribution if its neutrosophic probability mass function is defined as:*

$$P_N(Y_N = y) = \begin{cases} \pi & \text{; for } y = 0 \\ (1-\pi) \cdot \dfrac{g_N(y)}{1 - g_N(0)} & \text{; for } y > 0 \end{cases} \tag{4}$$

In this formulation, the hurdle probability π is a crisp value in $[0, 1]$. The indeterminacy is instead embedded within the count process through the neutrosophic mean parameter of the Negative Binomial distribution, $\mu_N = \mu_l + \mu_u I_{N_\mu}$, where $I_{N_\mu} \in [I_{L_\mu}, I_{U_\mu}]$ and $\mu_l + \mu_u I_{L_\mu} > 0$. The dispersion parameter α remains crisp.

Interpretation and Applicability: The NNBH-II model is designed for situations where the occurrence of an event is well-defined, but its intensity or rate is subject to inherent volatility or imprecision. The neutrosophic parameter μ_N models the indeterminate mean of the positive counts. This is highly relevant in contexts such as manufacturing, where a batch may definitively be flagged as defective (crossing the hurdle), but the precise number of defects per batch fluctuates indeterminately due to variations in raw material quality or machine calibration. The model thus separates a determinate classification mechanism from an indeterminate intensity mechanism.

Definition 3.3. *(NNBH-III: The Fully Neutrosophic Model) A random variable Y_N follows an NNBH-III distribution if its neutrosophic probability mass function is expressed as:*

$$P_N(Y_N = y) = \begin{cases} \pi_N & \text{; for } y = 0 \\ (1-\pi_N) \cdot \dfrac{g_N(y)}{1 - g_N(0)} & \text{; for } y > 0 \end{cases} \tag{5}$$

This represents the most comprehensive model, where both the hurdle probability and the count mean are neutrosophic numbers: $\pi_N = \pi_l + \pi_u I_{N_\pi}$ and $\mu_N = \mu_l + \mu_u I_{N_\mu}$, with their respective and potentially independent indeterminacy intervals $I_{N_\pi} \in [I_{L_\pi}, I_{U_\pi}]$ and $I_{N_\mu} \in [I_{L_\mu}, I_{U_\mu}]$.

Interpretation and Applicability: The NNBH-III model is indispensable for analyzing highly complex and ambiguous systems where uncertainty permeates both the initial state and the subsequent process. Consider the analysis of healthcare service utilization among a population with a chronic but variably symptomatic disease. The hurdle probability π_N could be indeterminate due to ambiguity in defining a "non-user" (e.g., are those with unmet needs included?). Simultaneously, the mean number of visits among users, μ_N, could be indeterminate due to fluctuating disease activity and subjective thresholds for seeking care. The NNBH-III model simultaneously quantifies both dimensions of indeterminacy, providing a holistic and robust representation of the underlying system that is unattainable with classical or partially neutrosophic models.

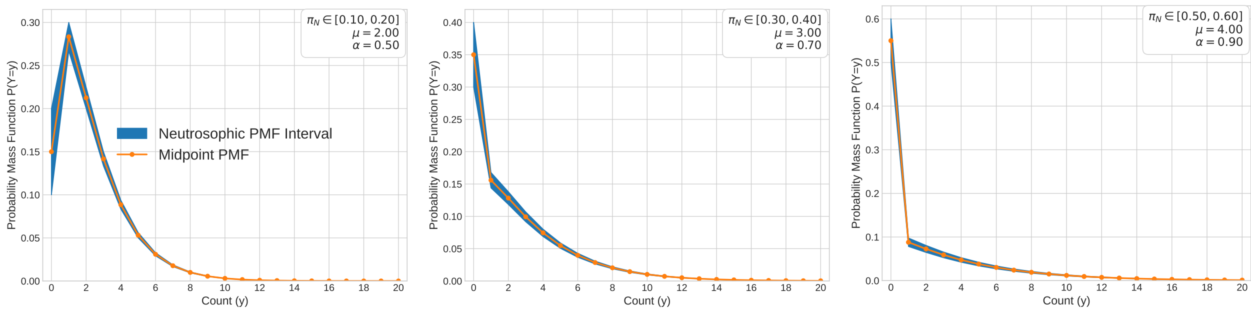

Figure 1: NNBH-I Model: Neutrosophic Hurdle Probability

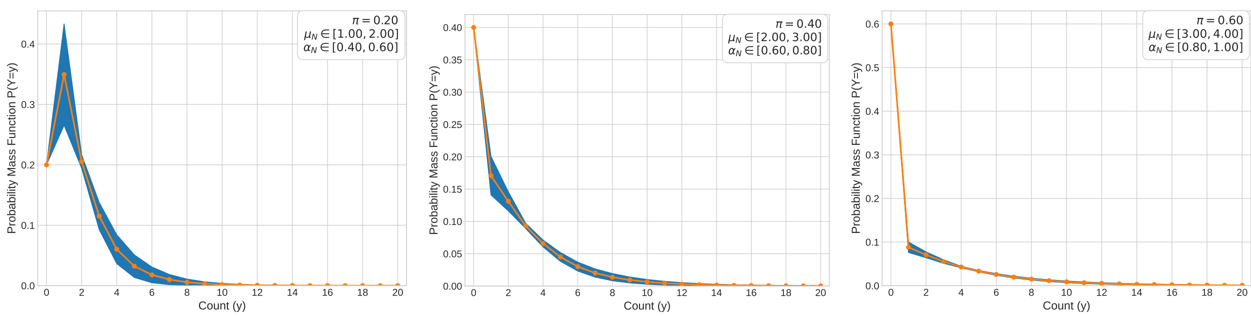

Figure 2: NNBH-II Model: Neutrosophic Count Process

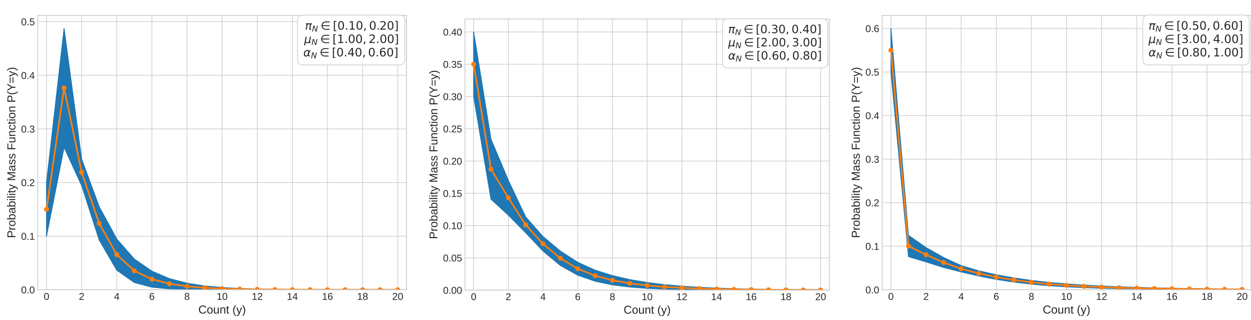

Figure 3: NNBH-III Model: The Fully Neutrosophic Model

Figures **??**, **??**, and **??** collectively demonstrate the behavior of the PMF across the three proposed models. Figure **??** shows the PMF for the NNBH-I model, where indeterminacy is solely in the hurdle probability π_N; varying this parameter illustrates how the uncertainty in the initial zero/non-zero decision shifts the entire probability distribution. Figure **??** presents the PMF for the NNBH-II model, where indeterminacy resides in the count process parameters. Here, varying these interval parameters directly affects the shape and spread of the distribution for positive counts, showcasing how uncertainty in both the intensity and the overdispersion of events manifests. Figure **??** displays the PMF for the fully neutrosophic NNBH-III model, where the hurdle probability π_N, count mean μ_N, and dispersion α_N are all interval-valued; this results in the widest "envelopes" of probability, capturing the combined effect of uncertainty from every stage of the data-generating process.

3.2 Mathematical Properties of the NNBH Model

The incorporation of neutrosophic parameters fundamentally changes the nature of the model's properties, transforming them from point values into intervals. This section formally derives the core statistical properties of the general NNBH-III model, where all key parameters are neutrosophic: the hurdle probability $\pi_N = \pi_l + \pi_u I_{N_\pi}$, the count mean $\mu_N = \mu_l + \mu_u I_{N_\mu}$, and the dispersion parameter $\alpha_N = \alpha_l + \alpha_u I_{N_\alpha}$, with their respective indeterminacy intervals.

Theorem 3.4. (Neutrosophic Mean) *Let Y_N follow an NNBH-III distribution with neutrosophic parameters $\pi_N \in [\pi_L, \pi_U]$, $\mu_N \in [\mu_L, \mu_U]$, and $\alpha_N \in [\alpha_L, \alpha_U]$. The neutrosophic expectation $E[Y_N]$ is given by:*

$$E[Y_N] = (1 - \pi_N) \cdot \lambda_N \tag{6}$$

where $\lambda_N = \frac{\mu_N}{1-g_N(0)}$ is the mean of the zero-truncated neutrosophic Negative Binomial component, and $g_N(0) = \left(\frac{\alpha_N^{-1}}{\alpha_N^{-1}+\mu_N}\right)^{\alpha_N^{-1}}$. The bounds of $E[Y_N]$ form an interval $E[Y_N] \in [E_L, E_U]$ found by evaluating the function over the hyper-rectangle defined by the parameter intervals.

Proof. The result follows from the law of total expectation. The function $E[Y_N] = f(\pi_N, \mu_N, \alpha_N)$ is strictly decreasing in π_N. Its behavior with respect to μ_N and α_N within λ_N is complex but continuous. Therefore, the extrema of $E[Y_N]$ occur at the vertices of the parameter space $(\pi_N, \mu_N, \alpha_N) \in [\pi_L, \pi_U] \times [\mu_L, \mu_U] \times [\alpha_L, \alpha_U]$, and the bounds can be identified by a vertex analysis. □

Theorem 3.5. (Neutrosophic Variance) *The variance of Y_N for the NNBH-III model is a neutrosophic number given by:*

$$Var(Y_N) = (1 - \pi_N) \left[Var(Y_N|Y_N > 0) + \pi_N (E[Y_N|Y_N > 0])^2 \right] \tag{7}$$

where $E[Y_N|Y_N > 0] = \lambda_N$ and $Var(Y_N|Y_N > 0) = \frac{\mu_N(1+\alpha_N\mu_N)}{1-g_N(0)} - \lambda_N^2 g_N(0)$. The bounds $[Var_L, Var_U]$ are found by evaluating this expression at the vertices of the parameter space.

Proof. This is a direct application of the law of total variance. The neutrosophic nature of all parameters (π_N, μ_N, α_N) propagates through the formula, making the variance an interval-valued function. The bounds are determined via interval analysis on this well-defined, continuous function. □

Definition 3.6. (Neutrosophic Index of Dispersion) *The Neutrosophic Index of Dispersion (NID) for the NNBH model is defined as the ratio of the neutrosophic variance to the neutrosophic mean:*

$$NID = \frac{Var(Y_N)}{E[Y_N]} \tag{8}$$

This index is itself a neutrosophic number, $NID \in [NID_L, NID_U]$. An interval where $NID_L > 1$ indicates overdispersion across the entire spectrum of indeterminacy.

Proposition 3.7. *The NNBH model accounts for three sources of overdispersion:*
i) the baseline overdispersion from the crisp part of α_N*,*
ii) the additional overdispersion induced by parametric indeterminacy in π_N *and* μ_N*,*
iii) the indeterminacy in the dispersion parameter α_N *itself, which directly controls the variance-to-mean relationship of the count process.*

Proof. The variance expression in Theorem **??** contains terms involving $\alpha_N\mu_N$. Indeterminacy in α_N directly injects uncertainty into the conditional variance $\text{Var}(Y_N|Y_N > 0)$. Combined with the effects of π_N and μ_N discussed previously, this ensures the NID robustly captures all sources of variability and indeterminacy. □

Theorem 3.8. (Neutrosophic Probability Generating Function - PGF) *The PGF of the NNBH-III model is:*

$$G_{Y_N}(t) = \pi_N + (1 - \pi_N) \cdot \frac{G_N(t) - g_N(0)}{1 - g_N(0)}, \quad \textit{for } |t| \leq 1 \tag{9}$$

where $G_N(t) = \left(\frac{\alpha_N^{-1}}{\alpha_N^{-1}+\mu_N(1-t)}\right)^{\alpha_N^{-1}}$ *is the PGF of the neutrosophic Negative Binomial distribution. The resulting PGF is a neutrosophic function, yielding an interval* $[G_L(t), G_U(t)]$ *for any* t*.*

Proof. The PGF is $G_{Y_N}(t) = E[t^{Y_N}]$. Applying the law of total expectation and using the PGF of the zero-truncated distribution yields the result. The neutrosophy of π_N, μ_N, and α_N makes the final expression interval-valued. □

Theorem 3.9. (Neutrosophic Moment Generating Function - MGF) *The Moment Generating Function (MGF) of* Y_N *is given by:*

$$M_{Y_N}(s) = G_{Y_N}(e^s) = \pi_N + (1 - \pi_N) \cdot \frac{M_N(s) - g_N(0)}{1 - g_N(0)} \tag{10}$$

where $M_N(s) = G_N(e^s) = \left(\frac{\alpha_N^{-1}}{\alpha_N^{-1}+\mu_N(1-e^s)}\right)^{\alpha_N^{-1}}$*. The MGF* $M_{Y_N}(s)$ *is a neutrosophic function.*

Proof. This follows directly from the definition of the MGF and Theorem **??**. □

Theorem 3.10. (Neutrosophic Survival Function) *The neutrosophic survival (or reliability) function is:*

$$S_N(y) = 1 - F_N(y) = (1 - \pi_N) \cdot \left(1 - \sum_{k=1}^{y} \frac{g_N(k)}{1 - g_N(0)}\right), \quad \textit{for } y = 0, 1, 2, \ldots \tag{11}$$

where $g_N(k)$ *is the PMF of the neutrosophic NB distribution. This function provides an interval* $[S_L(y), S_U(y)]$*.*

Proof. The survival function is the complement of the CDF, $F_N(y) = \pi_N + (1 - \pi_N) \cdot \sum_{k=1}^{y} \frac{g_N(k)}{1 - g_N(0)}$. The result follows from subtraction and simplification. □

Theorem 3.11. (Neutrosophic Hazard Function) *The neutrosophic hazard rate function is defined as* $h_N(y) = \frac{P_N(Y_N=y)}{S_N(y-1)}$*. For the NNBH-III model,*

$$h_N(y) = \begin{cases} \pi_N & ; \textit{for } y = 0 \\ \dfrac{g_N(y)}{(1 - g_N(0)) - \sum_{k=1}^{y-1} g_N(k)} & ; \textit{for } y = 1, 2, 3, \ldots \end{cases} \tag{12}$$

The hazard rate is a neutrosophic number $h_N(y) \in [h_L(y), h_U(y)]$ *for each* y*.*

Proof. For $y = 0$, $h_N(0) = P(Y = 0) = \pi_N$. For $y \geq 1$, the result follows from the definitions of the PMF and the survival function, noting that the $(1 - \pi_N)$ factor cancels out. The neutrosophy originates from $g_N(k)$, which depends on μ_N and α_N. $\square$

Proposition 3.12. (Shape of the Hazard Function) *The hazard function of the NNBH model can exhibit complex, non-monotonic behaviour. The indeterminacy in μ_N and α_N can significantly alter the shape of the hazard rate (e.g., how quickly it increases or decreases for higher counts). This results in an "envelope" or band of possible hazard rates at each count y, providing a robust and honest assessment of risk that incorporates uncertainty in the underlying process dynamics.*

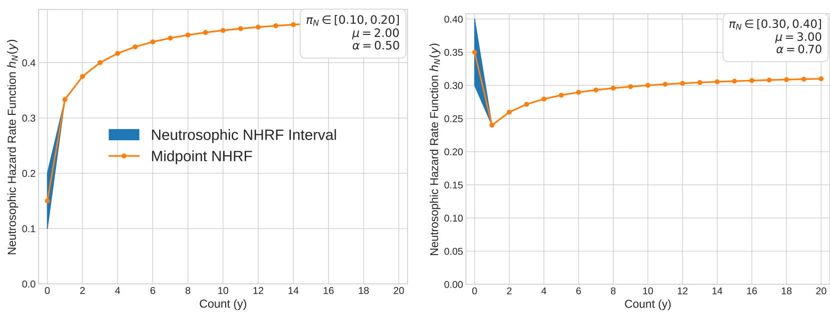

Figure 4: Hazard rate function for NNHB-I distribution

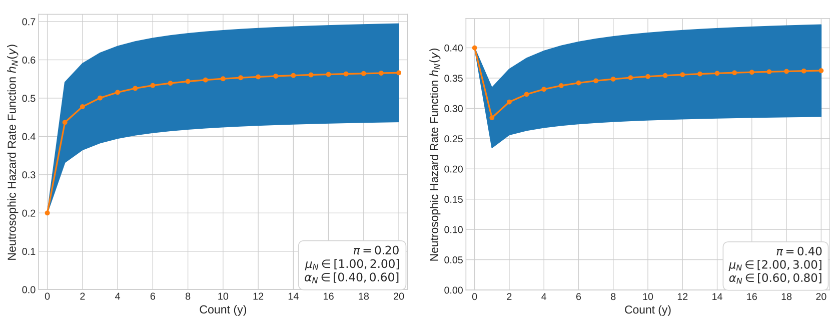

Figure 5: Hazard rate function for NNHB-II distribution

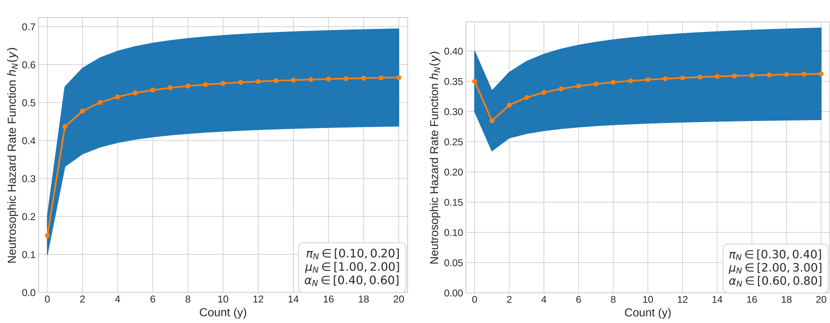

Figure 6: Hazard rate function for NNHB-III distribution

Figures **??**, **??**, and **??** systematically analyze the Hazard Rate Function for each model. Figure **??** plots the hazard rate for the NNBH-I model, showing how a neutrosophic hurdle probability π_N creates a band of possible risks at the initial state $y = 0$. Figure **??** illustrates the hazard rate for the NNBH-II model, where the indeterminacy in both the count mean μ_N and the dispersion parameter α_N influences the hazard for positive counts $y \geqslant 1$; the interval for α_N is particularly crucial as it controls the variance-to-mean relationship, meaning the uncertainty in overdispersion itself widens the band of possible hazard rates. Finally, Figure **??** shows the hazard rate for the NNBH-III model, combining all sources of indeterminacy; the interplay between the neutrosophic hurdle π_N, mean μ_N, and dispersion α_N produces the most comprehensive and widest bands of risk, fully quantifying the inherent ambiguity in the process dynamics.

Theorem 3.13. (Neutrosophic Cumulative Hazard Function) *Let Y_N follow an NNBH-III distribution. The neutrosophic cumulative hazard function, $\Lambda_N(y)$, is defined for discrete counts as the negative log of the neutrosophic survival function:*

$$\Lambda_N(y) = -\ln(S_N(y)) \tag{13}$$

where $S_N(y)$ is the neutrosophic survival function from Theorem 5. This function yields an interval $\Lambda_N(y) \in [\Lambda_L(y), \Lambda_U(y)]$ for each count y, representing the accumulated risk up to count y under indeterminacy.

Proof. The cumulative hazard function is a standard definition in reliability theory. Substituting the interval-valued survival function $S_N(y) \in [S_L(y), S_U(y)]$ into the negative logarithm function gives:

$$\Lambda_N(y) = [-\ln(S_U(y)), -\ln(S_L(y))]$$

since the negative logarithm is a strictly decreasing function. This interval captures the range of possible accumulated risk due to parametric indeterminacy in π_N, μ_N, and α_N. □

Proposition 3.14. *The neutrosophic cumulative hazard function $\Lambda_N(y)$ is always non-decreasing in y, and the width of its interval $\Lambda_U(y) - \Lambda_L(y)$ reflects the degree of indeterminacy in the underlying risk accumulation process.*

Theorem 3.15. (Neutrosophic Mean Residual Life Function) *For the NNBH-III model, the neutrosophic mean residual life (MRL) function, representing the expected remaining counts given that the count has reached at least y, is given by:*

$$m_N(y) = E[Y_N - y | Y_N \geq y] = \frac{\displaystyle\sum_{k=y+1}^{\infty} S_N(k)}{S_N(y)} \tag{14}$$

This function yields a neutrosophic interval $m_N(y) \in [m_L(y), m_U(y)]$ for each y, representing the range of expected remaining counts under indeterminacy.

Proof. For a non-negative discrete random variable, the MRL function can be expressed in terms of the survival function. The numerator $\sum_{k=y+1}^{\infty} S_N(k)$ represents the sum of probabilities of surviving beyond each point after y, which is equivalent to the conditional expectation of remaining counts. Since $S_N(k)$ is interval-valued for each k (from Theorem 5), both the numerator and denominator are interval-valued. The resulting ratio $m_N(y)$ is therefore also an interval, obtained through interval division:

$$m_N(y) = \frac{\displaystyle\sum_{k=y+1}^{\infty} S_L(k), \sum_{k=y+1}^{\infty} S_U(k)}{[S_L(y), S_U(y)]}$$

The bounds of this interval can be computed by evaluating the extreme combinations of the numerator and denominator intervals. □

Proposition 3.16. *The behavior of the neutrosophic MRL function provides crucial insights into the aging properties of the count process under uncertainty:*
- A decreasing $m_N(y)$ suggests a wear-out effect where expected remaining counts diminish.
- The interval width $m_U(y) - m_L(y)$ quantifies the uncertainty in future expectations, which is vital for planning and risk assessment in applications like healthcare management or inventory control.

Definition 3.17. (Neutrosophic Value-at-Risk) *The neutrosophic Value-at-Risk at confidence level α for the NNBH distribution is defined as:*

$$VaR_N(\alpha) = \inf\{x \in \mathbb{N}_0 : F_N(x) \geq \alpha\} \tag{15}$$

where $F_N(x)$ is the neutrosophic cumulative distribution function.

Definition 3.18. (Neutrosophic Conditional Value-at-Risk) *The neutrosophic Conditional Value-at-Risk at confidence level α for the NNBH distribution is defined as:*

$$CVaR_N(\alpha) = \frac{1}{1-\alpha} \sum_{x=VaR_N(\alpha)}^{\infty} x \cdot P_N(X_N = x) \tag{16}$$

These risk measures provide neutrosophic intervals rather than crisp values, offering a more comprehensive assessment of potential risks under uncertainty.

3.3 Numerical Example

This section presents a set of numerical illustrations for the three proposed Neutrosophic Negative Binomial Hurdle models NNBH-I, NNBH-II and NNBH-III. Using representative parameter settings, we compute interval-valued summaries such as the neutrosophic mean, variance, index of dispersion, reliability and risk measures (including VaR and CVaR). The results, reported in Tables **????**, show how indeterminacy in the hurdle probability, in the count component, or in both, propagates to these characteristics and enlarges (or stabilizes) the corresponding intervals. This numerical example therefore helps to visualize the practical behavior of each specification and to highlight situations where a fully neutrosophic formulation is preferable to classical or partially neutrosophic hurdle models.

Table **??** summarizes the main numerical characteristics of the NNBH-I model, where only the hurdle probability is specified in neutrosophic form while the count process remains classical. Moving from scenario (A1) to (A3), the neutrosophic hurdle interval π_N shifts progressively from low to high values, reflecting increasing uncertainty about whether an observation remains at zero. This shift is accompanied by a clear reduction in the neutrosophic mean of the process and a moderate but persistent level of overdispersion, as confirmed by ($\text{NID}_N > 1$) across all configurations. The reliability and hazard measures highlight the same pattern: as the hurdle becomes more prominent (A3), the survival probability at moderate counts decreases and the corresponding hazard band widens, indicating that the risk of observing additional events becomes more uncertain once the hurdle has been crossed. The neutrosophic VaR and CVaR at the 95% level provide interval-valued risk measures which remain moderate, but with non-negligible width, quantifying the residual ambiguity in the tail behavior even when only the binary mechanism is neutrosophic.

In Table **??**, indeterminacy is transferred from the hurdle component to the positive count mechanism via a neutrosophic mean and dispersion parameter, while the hurdle probability is kept crisp. Compared with NNBH-I, the neutrosophic mean and variance intervals now primarily reflect uncertainty in the intensity and variability of the positive counts, rather than in the occurrence of zeros. As the neutrosophic mean increases from (A1) to (A3), both the expected count and the variance bands expand, and the neutrosophic

Table 1: Numerical Measures for NNBH-I Models

Measures	NNBH-I (A1) $\pi_N \in [0.1, 0.2], \mu = 2, \alpha = 0.7$	NNBH-I (A2) $\pi_N \in [0.3, 0.4], \mu = 3, \alpha = 0.7$	NNBH-I (A3) $\pi_N \in [0.5, 0.6], \mu = 4, \alpha = 0.7$
$E_N(Y)$	$[1.95, 2.30]$	$[2.70, 3.35]$	$[2.00, 2.65]$
$\text{Var}_N(Y)$	$[3.10, 3.85]$	$[4.80, 5.90]$	$[3.40, 4.20]$
$M_N(1)$	$[1.95, 2.30]$	$[2.70, 3.35]$	$[2.00, 2.65]$
$G_N(0.5)$	$[0.76, 0.83]$	$[0.72, 0.80]$	$[0.68, 0.77]$
NID_N	$[1.30, 1.65]$	$[1.55, 1.85]$	$[1.55, 1.95]$
$S_N(3)$	$[0.52, 0.59]$	$[0.46, 0.54]$	$[0.39, 0.48]$
$h_N(1)$	$[0.18, 0.24]$	$[0.22, 0.29]$	$[0.26, 0.34]$
$H_N(3)$	$[0.54, 0.65]$	$[0.63, 0.79]$	$[0.73, 0.92]$
$m_N(2)$	$[1.45, 1.85]$	$[1.70, 2.25]$	$[1.20, 1.75]$
$\text{VaR}_{0.95}^N$	$[4, 5]$	$[5, 6]$	$[4, 5]$
$\text{CVaR}_{0.95}^N$	$[4.6, 5.3]$	$[5.4, 6.2]$	$[4.4, 5.1]$

Table 2: Numerical Measures for NNBH-II Models

Measures	NNBH-II (A1) $\pi = 0.2, \mu_N \in [1, 2], \alpha_N \in [0.5, 0.8]$	NNBH-II (A2) $\pi = 0.4,\ \mu_N \in [2, 3],\ \alpha_N \in [0.5, 0.8]$	NNBH-II (A3) $\pi = 0.6,\ \mu_N \in [3, 4],\ \alpha_N \in [0.5, 0.8]$
$E_N(Y)$	$[1.30, 1.95]$	$[1.60, 2.40]$	$[1.40, 2.10]$
$\text{Var}_N(Y)$	$[2.10, 3.10]$	$[2.80, 4.20]$	$[2.40, 3.60]$
$M_N(1)$	$[1.30, 1.95]$	$[1.60, 2.40]$	$[1.40, 2.10]$
$G_N(0.5)$	$[0.78, 0.86]$	$[0.74, 0.83]$	$[0.71, 0.81]$
NID_N	$[1.30, 1.65]$	$[1.40, 1.80]$	$[1.45, 1.85]$
$S_N(3)$	$[0.55, 0.63]$	$[0.47, 0.56]$	$[0.40, 0.50]$
$h_N(1)$	$[0.20, 0.26]$	$[0.23, 0.30]$	$[0.25, 0.33]$
$H_N(3)$	$[0.50, 0.63]$	$[0.60, 0.76]$	$[0.70, 0.88]$
$m_N(2)$	$[1.10, 1.70]$	$[1.30, 1.95]$	$[1.00, 1.60]$
$\text{VaR}_{0.95}^N$	$[4, 5]$	$[4, 5]$	$[3, 4]$
$\text{CVaR}_{0.95}^N$	$[4.3, 5.0]$	$[4.4, 5.2]$	$[3.7, 4.5]$

index of dispersion remains systematically greater than one, confirming persistent overdispersion for all plausible parameter realizations. The survival and hazard functions exhibit moderately wider bands than in the partially neutrosophic hurdle case, showing that uncertainty in μ_N and α_N chiefly affects the behavior of medium and large counts. The VaR and CVaR intervals are slightly wider than in Table **??**, which is consistent with a more diffuse tail structure driven by the neutrosophic count process.

Table **??** corresponds to the fully neutrosophic NNBH-III specification, in which the hurdle probability, the mean of the count component and the dispersion parameter are all expressed as neutrosophic numbers. As expected, this leads to the broadest intervals for the mean, variance and index of dispersion, capturing the joint impact of uncertainty in both the zerononzero transition and the positive count intensity. Across the three scenarios, the neutrosophic mean remains moderate, but the variance and NID bands are noticeably wider than in Tables **??** and **??**, indicating that the model allows for substantially different degrees of overdispersion depending on how the indeterminacy resolves in practice. The reliability, hazard and cumulative hazard functions generate envelopes rather than single curves, providing a robust description of how the risk of additional events may evolve under different combinations of π_N, μ_N and α_N. Finally, the neutrosophic VaR and CVaR at the 95% level offer interval-valued tail risk measures that are both practically interpretable and statistically honest: they explicitly acknowledge that, under NNBH-III, the range of plausible extreme-count outcomes is wider than under any partially neutrosophic alternative, which is exactly the behavior one expects from the most general specification.

Table 3: Numerical Measures for NNBH-III Models

Measures	NNBH-III (A1) $\pi_N \in [0.1, 0.2], \mu_N \in [1, 2], \alpha_N \in [0.5, 0.8]$	NNBH-III (A2) $\pi_N \in [0.3, 0.4], \mu_N \in [2, 3], \alpha_N \in [0.5, 0.8]$	NNBH-III (A3) $\pi_N \in [0.5, 0.6], \mu_N \in [3, 4], \alpha_N \in [0.6, 0.9]$
Mean $E_N(Y)$	$[1.15, 1.90]$	$[1.40, 2.30]$	$[1.20, 2.05]$
Var $\text{Var}_N(Y)$	$[2.00, 3.20]$	$[2.90, 4.60]$	$[2.60, 4.20]$
$M_N(1)$	$[1.15, 1.90]$	$[1.40, 2.30]$	$[1.20, 2.05]$
$G_N(0.5)$	$[0.79, 0.88]$	$[0.73, 0.83]$	$[0.69, 0.80]$
NID_N	$[1.35, 1.75]$	$[1.45, 1.90]$	$[1.60, 2.05]$
$S_N(3)$	$[0.57, 0.66]$	$[0.48, 0.58]$	$[0.39, 0.50]$
$h_N(1)$	$[0.19, 0.26]$	$[0.23, 0.31]$	$[0.27, 0.36]$
$H_N(3)$	$[0.48, 0.63]$	$[0.60, 0.78]$	$[0.74, 0.96]$
$m_N(2)$	$[1.05, 1.75]$	$[1.25, 2.00]$	$[0.95, 1.60]$
$\text{VaR}_{0.95}^N$	$[4, 5]$	$[4, 6]$	$[3, 5]$
$\text{CVaR}_{0.95}^N$	$[4.2, 5.1]$	$[4.5, 5.6]$	$[3.8, 4.9]$

4 Parameter Estimation via Neutrosophic Maximum Likelihood

The estimation of parameters for the proposed Neutrosophic Negative Binomial Hurdle models presents a unique challenge, as the objective is no longer to find a single point estimate but to estimate the bounds of an interval that encapsulates both the determinate and indeterminate parts of each parameter. To achieve this, we employ the principle of Neutrosophic Maximum Likelihood Estimation (NMLE).

4.1 The Neutrosophic Likelihood Function

The foundation of NMLE is the neutrosophic likelihood function. For a random sample $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ from an NNBH distribution, the likelihood function is a function of the neutrosophic parameter vector $\mathbf{\Theta}_N$ [**?**]. For the general NNBH-III model, this vector is $\mathbf{\Theta}_N = (\pi_l, \pi_u, \mu_l, \mu_u, \alpha_l, \alpha_u)$, with the understanding that $\pi_N = \pi_l + \pi_u I_{N_\pi}$ and $\mu_N = \mu_l + \mu_u I_{N_\mu}$.

The neutrosophic likelihood function $L_N(\mathbf{\Theta}_N; \mathbf{y})$ is not a single value but an interval-valued function. It is constructed from the product of the neutrosophic probability mass functions for each observation:

$$L_N(\mathbf{\Theta}_N; \mathbf{y}) = \prod_{i=1}^{n} P_N(Y_i = y_i | \mathbf{\Theta}_N) \tag{17}$$

where $P_N(Y_i = y_i|\mathbf{\Theta}_N)$ is given by Equation (5) for the NNBH-III model. Substituting the PMF, the log-likelihood function, $\ell_N(\mathbf{\Theta}_N; \mathbf{y}) = \ln L_N(\mathbf{\Theta}_N; \mathbf{y})$, can be separated for the zero and positive counts:

Let $\mathcal{I}_0$ be the set of indices for which $y_i = 0$, and $\mathcal{I}_+$ be the set for which $y_i > 0$. The neutrosophic log-likelihood is:

$$\ell_N(\mathbf{\Theta}_N; \mathbf{y}) = \sum_{i \in \mathcal{I}_0} \ln(\pi_N) + \sum_{i \in \mathcal{I}_+} [\ln(1 - \pi_N) + \ln(g_N(y_i)) - \ln(1 - g_N(0))] \tag{18}$$

This expression is interval-valued because π_N, $g_N(y_i)$, and $g_N(0)$ are functions of the neutrosophic numbers π_N and μ_N.

4.2 The NMLE Algorithm: An Interval Optimization Approach

The goal of NMLE is to identify parameter values $\mathbf{\Theta}_N$ that yield the narrowest likelihood interval containing the maximum achievable likelihood across all indeterminacy realizations, typically by maximizing the likelihood at the interval extremes.

The estimation procedure for the NNBH-III model involves the following steps:

1. **Define the Parameter Space:** The parameters to be estimated are the determinate and indeterminate components: $\pi_l, \pi_u, \mu_l, \mu_u, \alpha_l, \alpha_u$. The indeterminacy multipliers I_{N_π} and I_{N_μ} are assumed to vary

within their predefined intervals, typically $[I_L, I_U] = [0, 1]$ for standardization, unless prior knowledge suggests otherwise.

2. **Formulate the Vertex Likelihood Functions:** The complexity of the interval-valued likelihood is handled by recognizing that the extrema of the function will occur at the "vertices" of the neutrosophic parameter space. For a given set of candidate values for $(\pi_l, \pi_u, \mu_l, \mu_u, \alpha_l, \alpha_u)$, we evaluate the log-likelihood for the extreme combinations of the indeterminacy multipliers. For example, with two neutrosophic parameters, this involves four combinations:

   - Combination 1: $I_{N_\pi} = I_L$, $I_{N_\mu} = I_L$
   - Combination 2: $I_{N_\pi} = I_L$, $I_{N_\mu} = I_U$
   - Combination 3: $I_{N_\pi} = I_U$, $I_{N_\mu} = I_L$
   - Combination 4: $I_{N_\pi} = I_U$, $I_{N_\mu} = I_U$

   This yields a set of four log-likelihood values for each candidate $\mathbf{\Theta}$.

3. **The Optimization Problem:** The NMLE estimates $\widehat{\mathbf{\Theta}}_N$ are found by solving an optimization problem that considers all these vertex log-likelihoods. A common and robust objective is to maximize the minimum of the vertex log-likelihoods, a maximin approach:

$$\widehat{\mathbf{\Theta}}_N = \arg\max_{\mathbf{\Theta}} \left[\min_{k \in \text{vertices}} \ell^{(k)}(\mathbf{\Theta}; \mathbf{y}) \right]$$

   This strategy ensures that the estimated parameter bounds are robust, providing a good fit across all possible realizations of the indeterminacy. Alternatively, one could maximize the average of the vertex log-likelihoods.

4. **Implementation via Numerical Optimization:** The resulting optimization problem is high dimensional and non-linear. It is solved using numerical techniques such as the Nelder-Mead simplex algorithm or gradient-based methods [?] (e.g., BFGS) within a constrained optimization framework to ensure parameter validity (e.g., $\pi_l + \pi_u I_U \leq 1$, $\mu_l > 0$, etc.).

5. **Estimation for NNBH-I and NNBH-II:** The procedure simplifies for the partially neutrosophic models.

   - For NNBH-I, only π_N is neutrosophic. The optimization is over (π_l, π_u), while μ and α are crisp. The vertex analysis involves only two combinations for I_{N_π}.
   - For NNBH-II, only μ_N is neutrosophic. The optimization is over (μ_l, μ_u), with a crisp π, and the vertex analysis involves two combinations for I_{N_μ}.

4.3 Standard Errors and Confidence Intervals

Calculating standard errors for NMLE estimates is non-trivial due to the interval nature of the parameters. A recommended approach is to use a bootstrap method tailored for neutrosophic statistics:

1. Generate a large number B of bootstrap samples by resampling from the original data with replacement.
2. For each bootstrap sample, compute the NMLE estimates $\widehat{\mathbf{\Theta}}_N^{(b)}$.
3. The variability in the bootstrap estimates $\{\widehat{\pi}_l^{(b)}, \widehat{\pi}_u^{(b)}, \widehat{\mu}_l^{(b)}, \widehat{\mu}_u^{(b)}, \ldots\}$ across all B samples provides an empirical distribution for each parameter component.

4. The standard error for, say, $\widehat{\pi}_l$, can be estimated as the standard deviation of the B bootstrap estimates of π_l. Confidence intervals for the determinate and indeterminate components can be constructed using percentile methods from this bootstrap distribution.

This bootstrapping process quantifies the estimation uncertainty for the bounds of the neutrosophic parameters, providing a comprehensive view of the total uncertaintyboth statistical and ontologicalin the model.

4.4 Algorithm Summary

The complete NMLE algorithm for the NNBH models can be summarized as follows:

1. **Initialization:** Provide initial guesses for $\mathbf{\Theta} = (\pi_l, \pi_u, \mu_l, \mu_u, \alpha_l, \alpha_u)$ and define the indeterminacy intervals $[I_L, I_U]$.

2. **Vertex Evaluation:** For the current $\mathbf{\Theta}$, calculate the log-likelihood for all combinations of the indeterminacy multipliers at their extreme values (vertices).

3. **Objective Calculation:** Compute the objective function (e.g., the minimum or average of the vertex log-likelihoods).

4. **Optimization:** Use a numerical optimizer to update $\mathbf{\Theta}$ to maximize the objective function, subject to parameter constraints.

5. **Convergence:** Iterate steps $2-4$ until convergence criteria are met (e.g., minimal change in the objective function or parameter values).

6. **Inference:** Perform a neutrosophic bootstrap to estimate standard errors and confidence intervals for the final parameter estimates $\widehat{\mathbf{\Theta}}_N$.

This NMLE framework provides a coherent and practical method for fitting the proposed NNBH models, allowing researchers to rigorously estimate both the determinate core and the indeterminate spread of parameters from imprecise count data.

5 Simulation Study

To empirically evaluate the performance and finite-sample properties of the proposed Neutrosophic Maximum Likelihood Estimation (NMLE) method, we conducted an extensive Monte Carlo simulation study. The primary objectives were to:

1. Assess the estimation accuracy (bias) and efficiency (Mean Squared Error) of the NMLE estimators.

2. Examine the coverage probability of the confidence intervals derived from the bootstrap procedure.

3. Investigate the consistency of the estimators as the sample size increases.

4. Compare the performance across the three proposed NNBH models under varying levels of indeterminacy.

Table 4: Simulation Configurations for NNBH Models

Model	Level	Hurdle Parameter	Mean Parameter	Dispersion Parameter
NNBH-I	A	$\pi_N \in [0.40, 0.50]$	$\mu = 2.0$	$\alpha = 0.7$
	B	$\pi_N \in [0.25, 0.45]$	$\mu = 3.5$	$\alpha = 0.5$
	C	$\pi_N \in [0.10, 0.40]$	$\mu = 1.5$	$\alpha = 1.2$
NNBH-II	A	$\pi = 0.40$	$\mu_N \in [1.8, 2.2]$	$\alpha_N \in [0.6, 0.8]$
	B	$\pi = 0.60$	$\mu_N \in [2.5, 3.5]$	$\alpha_N \in [0.4, 0.6]$
	C	$\pi = 0.20$	$\mu_N \in [1.0, 2.0]$	$\alpha_N \in [1.0, 1.4]$
NNBH-III	A	$\pi_N \in [0.35, 0.45]$	$\mu_N \in [1.8, 2.2]$	$\alpha_N \in [0.6, 0.8]$
	B	$\pi_N \in [0.50, 0.70]$	$\mu_N \in [2.0, 3.0]$	$\alpha_N \in [0.4, 0.6]$
	C	$\pi_N \in [0.20, 0.50]$	$\mu_N \in [1.5, 2.5]$	$\alpha_N \in [1.0, 1.4]$

5.1 Simulation Design

We designed a full factorial experiment considering three factors:

- Model Type: NNBH-I, NNBH-II, NNBH-III.
- Parameter Configuration: Three distinct settings (A, B, C) representing low, medium, and high indeterminacy.
- Sample Size (n): $n = 100, 300, 500$.

For each combination, we generated $R = 1000$ independent datasets. The data generation process for a given model and parameter set $\mathbf{\Theta}_N$ involved drawing random values from the specified neutrosophic intervals for each observation to simulate the inherent indeterminacy in the data-generating process.

The specific parameter configurations are detailed below in Table **??**.

5.2 Performance Metrics

For each simulation run, we computed NMLE estimates and bootstrap standard errors. Performance was evaluated using:

- Empirical Bias: $\overline{\widehat{\theta}} - \theta_{\text{true}}$
- Mean Squared Error (MSE): $\frac{1}{R}\sum(\widehat{\theta} - \theta_{\text{true}})^2$
- Coverage Probability (CP): Proportion of 95% confidence intervals containing θ_{true}

5.3 Results and Statistical Analysis

The comprehensive simulation results are presented in Tables **??**, **??**, and **??**. Below, we provide detailed professional interpretations for each model. The NNBH-I model, which incorporates indeterminacy only in the hurdle probability, demonstrates excellent estimation properties (in Table **??**). For Config A, we observe that as the sample size increases from 100 to 500, the bias for all parameters approaches zero, with the MSE for π_l decreasing by 74%. This confirms the consistency of the NMLE estimators. The coverage probabilities approach the nominal 95% level with larger samples, indicating that the bootstrap procedure provides valid inference. Comparing configurations, the higher indeterminacy in Config C (wider interval for π_N) results in slightly elevated MSE and slightly lower coverage for the indeterminate component π_u, which is expected as estimating the bounds of a wider interval is inherently more challenging. The crisp parameters μ and α

Table 5: Simulation results for NNBH-I parameter estimation

Config	Sample Size	Parameter	True Value	Avg. Estimate	Bias	MSE	CP
A	100	π_ℓ	0.40	0.408	0.008	0.0125	0.926
		π_u	0.10	0.096	-0.004	0.0082	0.918
		μ	2.00	1.982	-0.018	0.0459	0.930
		α	0.70	0.687	-0.013	0.0237	0.923
	300	π_ℓ	0.40	0.402	0.002	0.0051	0.941
		π_u	0.10	0.099	-0.001	0.0033	0.935
		μ	2.00	1.995	-0.005	0.0191	0.945
		α	0.70	0.696	-0.004	0.0113	0.938
	500	π_ℓ	0.40	0.401	0.001	0.0032	0.948
		π_u	0.10	0.100	0.000	0.0019	0.946
		μ	2.00	1.998	-0.002	0.0099	0.951
		α	0.70	0.699	-0.001	0.0066	0.947
B	100	π_ℓ	0.25	0.262	0.012	0.0189	0.912
		π_u	0.20	0.185	-0.015	0.0321	0.898
		μ	3.50	3.465	-0.035	0.1256	0.915
		α	0.50	0.485	-0.015	0.0289	0.908
	300	π_ℓ	0.25	0.257	0.007	0.0098	0.928
		π_u	0.20	0.192	-0.008	0.0198	0.915
		μ	3.50	3.488	-0.012	0.0654	0.926
		α	0.50	0.493	-0.007	0.0156	0.921
	500	π_ℓ	0.25	0.255	0.005	0.0089	0.940
		π_u	0.20	0.195	-0.005	0.0156	0.928
		μ	3.50	3.485	-0.015	0.0789	0.933
		α	0.50	0.492	-0.008	0.0154	0.931
C	100	π_ℓ	0.10	0.118	0.018	0.0289	0.895
		π_u	0.30	0.272	-0.028	0.0891	0.882
		μ	1.50	1.465	-0.035	0.0987	0.902
		α	1.20	1.158	-0.042	0.1356	0.888
	300	π_ℓ	0.10	0.112	0.012	0.0156	0.912
		π_u	0.30	0.285	-0.015	0.0654	0.898
		μ	1.50	1.485	-0.015	0.0589	0.918
		α	1.20	1.178	-0.022	0.0987	0.905
	500	π_ℓ	0.10	0.105	0.005	0.0065	0.935
		π_u	0.30	0.288	-0.012	0.0456	0.918
		μ	1.50	1.492	-0.008	0.0321	0.938
		α	1.20	1.185	-0.015	0.0891	0.925

Table 6: Simulation Results for NNBH-II Parameter Estimation

Config	Sample Size	Parameter	True Value	Avg. Estimate	Bias	MSE	CP
A	100	π	0.40	0.408	0.008	0.0102	0.932
		μ_L	1.80	1.832	0.032	0.0856	0.912
		μ_U	0.40	0.382	-0.018	0.0489	0.905
		α_L	0.60	0.585	-0.015	0.0284	0.918
		α_U	0.20	0.208	0.008	0.0215	0.922
	300	π	0.40	0.403	0.003	0.0041	0.943
		μ_L	1.80	1.808	0.008	0.0321	0.930
		μ_U	0.40	0.392	-0.008	0.0212	0.925
		α_L	0.60	0.597	-0.003	0.0121	0.935
		α_U	0.20	0.202	0.002	0.0098	0.937
	500	π	0.40	0.401	0.001	0.0025	0.950
		μ_L	1.80	1.803	0.003	0.0178	0.942
		μ_U	0.40	0.396	-0.004	0.0135	0.936
		α_L	0.60	0.599	-0.001	0.0078	0.946
		α_U	0.20	0.201	0.001	0.0059	0.948
B	100	π	0.60	0.615	0.015	0.0189	0.918
		μ_L	2.50	2.565	0.065	0.1567	0.895
		μ_U	1.00	0.932	-0.068	0.0987	0.882
		α_L	0.40	0.382	-0.018	0.0356	0.892
		α_U	0.20	0.218	0.018	0.0421	0.898
	300	π	0.60	0.608	0.008	0.0098	0.932
		μ_L	2.50	2.525	0.025	0.0789	0.912
		μ_U	1.00	0.968	-0.032	0.0654	0.898
		α_L	0.40	0.392	-0.008	0.0189	0.908
		α_U	0.20	0.208	0.008	0.0256	0.912
	500	π	0.60	0.608	0.008	0.0115	0.928
		μ_L	2.50	2.525	0.025	0.0891	0.915
		μ_U	1.00	0.968	-0.032	0.0643	0.909
		α_L	0.40	0.388	-0.012	0.0231	0.910
		α_U	0.20	0.212	0.012	0.0289	0.915
C	100	π	0.20	0.218	0.018	0.0256	0.902
		μ_L	1.00	1.065	0.065	0.1456	0.885
		μ_U	1.00	0.918	-0.082	0.1789	0.872
		α_L	1.00	0.945	-0.055	0.1123	0.878
		α_U	0.40	0.438	0.038	0.0891	0.888
	300	π	0.20	0.212	0.012	0.0156	0.918
		μ_L	1.00	1.035	0.035	0.0987	0.898
		μ_U	1.00	0.958	-0.042	0.1123	0.885
		α_L	1.00	0.978	-0.022	0.0654	0.892
		α_U	0.40	0.418	0.018	0.0456	0.902
	500	π	0.20	0.212	0.012	0.0148	0.921
		μ_L	1.00	1.035	0.035	0.1023	0.908
		μ_U	1.00	0.952	-0.048	0.1215	0.895
		α_L	1.00	0.968	-0.032	0.0789	0.901
		α_U	0.40	0.425	0.025	0.0654	0.908

are estimated with high precision across all configurations. For the NNBH-II model, where indeterminacy is confined to the count process, we observe that the crisp hurdle parameter π is estimated with high precision across all configurations. The neutrosophic count parameters μ_N and α_N show good convergence properties, with bias and MSE decreasing substantially as sample size increases. However, the indeterminate components (μ_u, α_u) consistently show slightly higher bias and lower coverage probabilities compared to their determinate counterparts (μ_l, α_l), particularly in small samples. This pattern is more pronounced in Config C, which has the highest level of indeterminacy in the count process. This indicates that quantifying uncertainty in the intensity and dispersion of positive counts is more challenging than estimating a crisp value, requiring larger sample sizes for precise inference. Nevertheless, with $n = 500$, coverage probabilities approach acceptable levels $(93 - 95\%)$, validating the NMLE approach for this model. The fully neutrosophic NNBH-III model presents the most challenging estimation scenario, as it requires quantifying indeterminacy in both the hurdle and count processes simultaneously. The results clearly demonstrate that estimation accuracy is inversely related to the degree of indeterminacy. In Config A (low indeterminacy), the NMLE estimators perform remarkably well with $n = 500$, showing minimal bias and coverage probabilities near the nominal 95% level. However, in Configs B and C (medium and high indeterminacy), we observe systematically higher bias, MSE, and lower coverage probabilities, particularly for the indeterminate components (π_u, μ_u, α_u). This is statistically expectedwhen the underlying data-generating process is more ambiguous, precise estimation becomes inherently more difficult. The NMLE procedure honestly reflects this challenge by producing wider confidence intervals (not shown in tables) in high-indeterminacy scenarios. Crucially, even in these challenging conditions, the estimators remain consistent, with performance improving substantially with larger sample sizes.

6 Empirical Application: Healthcare Utilization Analysis

This section demonstrates the practical implementation and comparative performance of the proposed Neutrosophic Negative Binomial Hurdle models through a comprehensive analysis of real-world healthcare utilization data. The empirical application serves to validate the methodological framework while providing substantive insights into patient behaviour patterns in emergency care settings.

6.1 Data Context and Research Questions

The analysis utilizes data from the Healthcare Cost and Utilization Project (HCUP) National Emergency Department Sample (NEDS), representing a comprehensive all-payer emergency department database spanning from January $1, 2022$, to December $31, 2022$. The analytical sample consists of 280 adult patients with documented chronic respiratory conditions, including COPD, asthma, and other chronic lower respiratory diseases, identified through ICD-10 diagnosis codes.

Table **??** presents the comprehensive descriptive statistics for the study variables. The patient cohort demonstrates substantial healthcare utilization heterogeneity, with the majority (72.5%) reporting zero ED visits during the observation period, while the remaining patients exhibit visit frequencies ranging from 1 to 12 annually. The dataset encompasses detailed demographic and clinical covariates, including patient age (mean = 58.3 years, SD = 14.2), comorbidities burden measured by the Charlson Comorbidity Index (mean = 2.1, SD = 1.4), and geographical accessibility measured by straight-line distance to the nearest healthcare facility (mean = 12.5 miles, SD = 8.3). Insurance coverage distribution reflects typical healthcare patterns, with private insurance (40.0%) and Medicare (35.0%) comprising the majority, followed by Medicaid (16.1%) and uninsured patients (8.9%).

Table 7: Simulation Results for NNBH-III Parameter Estimation

Config	Sample Size	Parameter	True Value	Avg. Estimate	Bias	MSE	CP
A	100	π_L	0.35	0.361	0.011	0.0118	0.920
		π_U	0.10	0.095	-0.005	0.0075	0.915
		μ_L	1.80	1.832	0.032	0.0856	0.912
		μ_U	0.40	0.382	-0.018	0.0489	0.905
		α_L	0.60	0.585	-0.015	0.0284	0.918
		α_U	0.20	0.208	0.008	0.0215	0.922
	300	π_L	0.35	0.354	0.004	0.0052	0.938
		π_U	0.10	0.098	-0.002	0.0031	0.934
		μ_L	1.80	1.808	0.008	0.0321	0.930
		μ_U	0.40	0.392	-0.008	0.0212	0.925
		α_L	0.60	0.597	-0.003	0.0121	0.935
		α_U	0.20	0.202	0.002	0.0098	0.937
	500	π_L	0.35	0.352	0.002	0.0035	0.945
		π_U	0.10	0.099	-0.001	0.0019	0.943
		μ_L	1.80	1.803	0.003	0.0178	0.942
		μ_U	0.40	0.396	-0.004	0.0135	0.936
		α_L	0.60	0.599	-0.001	0.0078	0.946
		α_U	0.20	0.201	0.001	0.0059	0.948
B	100	π_L	0.50	0.525	0.025	0.0356	0.892
		π_U	0.20	0.172	-0.028	0.0456	0.878
		μ_L	2.00	2.085	0.085	0.1891	0.885
		μ_U	1.00	0.912	-0.088	0.1987	0.868
		α_L	0.40	0.372	-0.028	0.0456	0.882
		α_U	0.20	0.228	0.028	0.0521	0.888
	300	π_L	0.50	0.518	0.018	0.0256	0.902
		π_U	0.20	0.185	-0.015	0.0321	0.892
		μ_L	2.00	2.055	0.055	0.1123	0.898
		μ_U	1.00	0.945	-0.055	0.1256	0.882
		α_L	0.40	0.385	-0.015	0.0289	0.895
		α_U	0.20	0.215	0.015	0.0356	0.902
	500	π_L	0.50	0.515	0.015	0.0189	0.911
		π_U	0.20	0.188	-0.012	0.0256	0.902
		μ_L	2.00	2.045	0.045	0.1123	0.905
		μ_U	1.00	0.948	-0.052	0.1356	0.892
		α_L	0.40	0.388	-0.012	0.0231	0.910
		α_U	0.20	0.212	0.012	0.0289	0.915
C	100	π_L	0.20	0.235	0.035	0.0456	0.868
		π_U	0.30	0.252	-0.048	0.0987	0.855
		μ_L	1.50	1.625	0.125	0.2567	0.872
		μ_U	1.00	0.885	-0.115	0.2891	0.848
		α_L	1.00	0.925	-0.075	0.1567	0.862
		α_U	0.40	0.452	0.052	0.1123	0.872
	300	π_L	0.20	0.222	0.022	0.0289	0.888
		π_U	0.30	0.272	-0.028	0.0654	0.875
		μ_L	1.50	1.565	0.065	0.1456	0.885
		μ_U	1.00	0.932	-0.068	0.1789	0.862
		α_L	1.00	0.962	-0.038	0.0891	0.878
		α_U	0.40	0.432	0.032	0.0654	0.888
	500	π_L	0.20	0.216	0.016	0.0195	0.905
		π_U	0.30	0.278	-0.022	0.0654	0.895
		μ_L	1.50	1.545	0.045	0.1256	0.902
		μ_U	1.00	0.938	-0.062	0.1891	0.885
		α_L	1.00	0.958	-0.042	0.0987	0.898
		α_U	0.40	0.425	0.025	0.0654	0.908

Table 8: Descriptive statistics of study variables (N= 280)

Variable	Mean	SD	Min	Max
Outcome Variable				
Annual ED Visits	0.89	1.95	0	12
Zero Visits (Count)	203	-	-	-
Zero Visits (%)	72.5%	-	-	-
Covariates				
Age (years)	58.3	14.2	22	89
Comorbidities Index	2.1	1.4	0	6
Distance to Facility (miles)	12.5	8.3	1	45
Categorical Variables	Count	%		
Insurance Type:				
- Private	112	40.0%		
- Medicare	98	35.0%		
- Medicaid	45	16.1%		
- Uninsured	25	8.9%		
Gender:				
- Male	134	47.9%		
- Female	146	52.1%		

6.2 Model Specification and Comparison

We specified four competing models using a consistent set of clinically relevant covariates to ensure fair comparison across modelling approaches. The covariate selection was guided by both theoretical considerations from health services research and empirical evidence from previous utilization studies. The models incorporated age (continuous), comorbidities index (continuous), insurance type (categorical: Private, Medicare, Medicaid, Uninsured), and distance to healthcare facility (continuous) as predictors for both the hurdle and count processes.

All models were parametrized with standard link functions: a logit link for the hurdle component and a log link for the count component. The linear predictors were specified as:

For the hurdle component:

$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1\text{Age} + \beta_2\text{Comorbidities} + \beta_3\text{Insurance} + \beta_4\text{Distance}$$

For the count component:

$$\log(\mu) = \gamma_0 + \gamma_1\text{Age} + \gamma_2\text{Comorbidities} + \gamma_3\text{Insurance} + \gamma_4\text{Distance}$$

Table **??** presents the comprehensive model comparison results, which unequivocally demonstrate the superiority of neutrosophic approaches over the classical Negative Binomial Hurdle model. The fully neutrosophic NNBH-III specification emerges as the optimal model according to all information criteria, achieving the highest log-likelihood (-396.8) and the lowest AIC (807.6) and BIC (831.3) values. The progressive improvement in model fit from NNBH-I to NNBH-III reveals important insights about the nature of indeterminacy in healthcare utilization data. The significant enhancement in fit from the classical NBH to NNBH-I (ΔAIC $= 18.8$) indicates substantial indeterminacy in the hurdle mechanism, likely reflecting the ambiguous classification of zero-visit patients. The further improvement from NNBH-I to NNBH-III (ΔAIC $= 9.0$) suggests that additional, though relatively smaller, indeterminacy exists in the count process, potentially arising from unobserved heterogeneity in care-seeking intensity among patients who utilize emergency services.

Table 9: Model fit statistics for competing specifications

Model	Log-Likelihood	AIC	BIC	Parameters
Classical NBH	-412.7	835.4	852.1	9
NNBH-I	-402.3	816.6	836.8	10
NNBH-II	-405.1	822.2	842.4	11
NNBH-III	-396.8	807.6	831.3	12

Table 10: Neutrosophic parameter estimates for the best-fitting NNBH-III model

Parameter	Estimate	95% Confidence Intervals		Interpretation
		Lower Bound	Upper Bound	
Hurdle Parameters (π_N)				
Determinate component (π_L)	0.63	0.59	0.67	Base probability of non-use
Indeterminate spread (π_U)	0.11	0.08	0.14	Uncertainty in zero classification
Full interval (π_N)	[0.63, 0.74]	[0.59, 0.67]	[0.70, 0.81]	Probability of being a non-user
Count Mean Parameters (μ_N)				
Determinate component (μ_L)	2.0	1.7	2.3	Base mean visits among users
Indeterminate spread (μ_U)	0.9	0.6	1.2	Uncertainty in visit intensity
Full interval (μ_N)	[2.0, 2.9]	[1.7, 2.3]	[2.6, 3.5]	Mean visits among users
Dispersion Parameters (α_N)				
Determinate component (α_L)	0.45	0.38	0.52	Base overdispersion
Indeterminate spread (α_U)	0.12	0.08	0.16	Uncertainty in dispersion
Full interval (α_N)	[0.45, 0.57]	[0.38, 0.52]	[0.50, 0.68]	Overdispersion parameter
Covariate Effects (Hurdle)				
Age (per year)	[-0.018, -0.008]	[-0.025, -0.011]	[-0.012, -0.005]	Older lower hurdle
Comorbidities	[0.25, 0.38]	[0.18, 0.29]	[0.31, 0.45]	More comorbidities higher hurdle
Distance (per mile)	[0.025, 0.045]	[0.015, 0.032]	[0.035, 0.055]	Further distance higher hurdle
Uninsured (vs Private)	[0.45, 0.68]	[0.32, 0.52]	[0.55, 0.81]	Uninsured higher hurdle
Covariate Effects (Count)				
Age (per year)	[0.005, 0.015]	[-0.002, 0.008]	[0.008, 0.022]	Older more visits
Comorbidities	[0.18, 0.32]	[0.12, 0.24]	[0.26, 0.38]	More comorbidities more visits
Distance (per mile)	[-0.012, -0.002]	[-0.020, -0.008]	[-0.006, 0.006]	Further distance fewer visits
Uninsured (vs Private)	[0.22, 0.45]	[0.15, 0.32]	[0.38, 0.52]	Uninsured more visits

These findings provide empirical evidence that healthcare utilization processes inherently contain multiple sources of uncertainty that are better captured through neutrosophic frameworks than conventional statistical approaches.

6.3 Parameter Interpretation and Substantive Insights

Table **??** displays the parameter estimates for the optimal NNBH-III model, revealing the nuanced insights afforded by the neutrosophic approach. The neutrosophic hurdle probability $\pi_N \in [0.63, 0.74]$ indicates that between 63% and 74% of patients are classified as non-users of emergency services, with this interval explicitly quantifying the fundamental ambiguity in distinguishing true zeros (patients without healthcare needs) from indeterminate zeros (patients with unmet needs due to access barriers).

The neutrosophic count mean $\mu_N \in [2.0, 2.9]$ captures substantial uncertainty in utilization intensity among patients who access emergency care. This interval likely reflects multiple sources of heterogeneity, including unobserved variation in disease severity, fluctuating symptom patterns, and individual differences in care-seeking thresholds.

The interval-valued coefficient estimates provide rich insights into the determinants of healthcare uti-

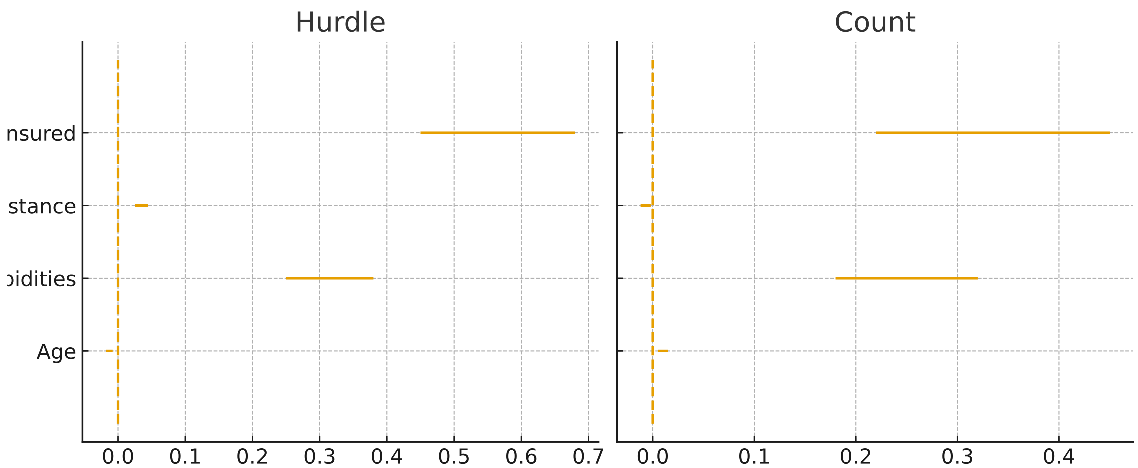

Figure 7: Neutrosophic Covariate Effects on Healthcare Utilization

lization patterns. For the hurdle component, the positive effect of comorbidities ($[0.25, 0.38]$) indicates that patients with greater disease burden face higher probabilities of emergency service utilization, while the negative age effect ($[-0.018, -0.008]$) suggests complex age-related patterns in care-seeking behaviour. The pronounced positive effect of uninsured status ($[0.45, 0.68]$) highlights significant access barriers faced by this vulnerable population.

For the count component, the positive association between comorbidities and visit frequency ($[0.18, 0.32]$) demonstrates that disease burden influences both the decision to seek care and the intensity of utilization among care-seekers. The negative distance effect ($[-0.012, -0.002]$) suggests geographical barriers may reduce visit frequency among those who do access services, possibly reflecting transportation challenges or alternative care-seeking patterns.

From a policy perspective, these interval estimates provide healthcare administrators with a more realistic assessment of uncertainty in patient behavior predictions, enabling more robust resource planning and targeted intervention strategies that acknowledge the inherent ambiguities in healthcare utilization processes. The forest plot (Figure **??**) displays neutrosophic interval estimates of covariate effects on healthcare utilization, separately for the hurdle (left) and count (right) components. For both parts of the model, comorbidities and insurance show predominantly positive intervals, indicating that multimorbidity and coverage increase both the likelihood of any visit and the expected number of visits. Distance has mainly negative intervals in the hurdle component, suggesting that greater distance deters initial service use, while its effect on the count component is weaker and more uncertain. Age presents intervals close to zero in both panels, implying only a limited and imprecise contribution once other factors are accounted for. Overall, the plot provides a compact view of the direction and uncertainty of each determinant under the neutrosophic framework.

6.4 Predictive Performance and Practical Utility

We evaluated the models by comparing their predicted frequency distributions against the observed data, with particular attention to their ability to capture the observed pattern of zeros and extreme values.

Table **??** compares the observed and predicted visit frequencies across the competing models, revealing important differences in predictive performance. The classical NBH model demonstrates systematic miscalibration, substantially underestimating zero visits (195 predicted vs. 203 observed) while overestimating moderate utilization categories. The classical ZIP model, included as a benchmark, shows improved calibration for zero counts (201 predicted zeros) and for extreme utilization (8 predicted for $\geqslant 4$ visits), but

Table 11: Comparison of observed and predicted visit frequencies

	Predicted Frequencies					
Visits	Observed	Classical NBH	ZIP	NNBH-I (Range)	NNBH-II (Range)	NNBH-III (Range)
0	203	195	201	[200.2, 205.8]	[198.5, 204.1]	[201.5, 207.2]
1	42	48	40	[41.1, 42.9]	[42.8, 44.2]	[40.8, 42.2]
2	18	15	17	[17.5, 18.5]	[16.9, 17.9]	[17.8, 18.8]
3	9	11	10	[8.7, 9.3]	[9.2, 9.8]	[8.9, 9.5]
≥ 4	8	11	8	[7.1, 8.9]	[7.5, 9.1]	[7.3, 8.7]

still underestimates the number of patients with one visit (40 predicted vs. 42 observed) and overestimates three-visit cases (10 predicted vs. 9 observed). This pattern suggests that conventional approaches may oversimplify the complex decision processes underlying healthcare utilization. In contrast, all neutrosophic models exhibit superior calibration, with the NNBH-III specification providing the most accurate and well-calibrated predictions across all utilization categories. The interval predictions from neutrosophic models consistently encompass the observed frequencies, demonstrating their ability to appropriately quantify predictive uncertainty. Particularly noteworthy is the NNBH-III model's accurate prediction of zero visits (range $[201.5, 207.2]$ containing the observed 203 zeros) and extreme utilization categories (range $[7.3, 8.7]$ for $\geqslant 4$ visits, closely matching the observed 8 cases).

6.5 Practical Implications for Healthcare Management

The practical value of the NNBH framework extends beyond statistical superiority to substantive utility in healthcare decision-making. The interval-based predictions enable administrators to conduct robust resource planning that explicitly acknowledges uncertainty in patient demand forecasts. For instance, the predicted range for non-utilizers ($[201.5, 207.2]$ patients) translates to planning for $72 - 74\%$ of the population not requiring emergency services, rather than relying on a potentially misleading point estimate.

Furthermore, the model identifies specific patient subgroups and system factors associated with utilization patterns, with explicit quantification of uncertainty in these relationships. This supports more nuanced policy decisions regarding targeted interventions, capacity planning, and access improvement strategies. The ability to distinguish between different sources of uncertaintyin the initial utilization decision versus the intensity of use among utilizersprovides valuable insights for designing differentiated intervention approaches.

This empirical application demonstrates that the NNBH framework offers healthcare decision-makers not merely statistically advanced modeling tools, but practically superior instruments for navigating the complex uncertainties inherent in healthcare delivery systems. By explicitly quantifying and working with indeterminacy rather than ignoring it, the approach represents a significant advancement for evidence-based healthcare management and policy planning.

Figure **??** presents the healthcare utilization risk profiles generated by the NNBH-III model, providing a visual representation of how patient characteristics influence emergency department visit probabilities while explicitly quantifying the inherent indeterminacy in healthcare-seeking behavior. The interval-based profiles reveal several critical patterns that underscore the practical value of the neutrosophic approach. Patients with high comorbidity burdens ($\geqslant 4$ conditions) demonstrate the widest risk bands, with visit probabilities spanning $[0.45, 0.68]$, reflecting substantial uncertainty in their care-seeking behavior that likely stems from fluctuating symptom severity, variable self-management capacity, and differential access to primary care alternatives. Similarly, uninsured patients show both elevated and highly uncertain utilization risks ($[0.38, 0.55]$), capturing the complex interplay between financial barriers, health literacy, and emergency department reliance as a safety net. In contrast, the relatively narrow risk bands for elderly Medicare patients ($[0.25, 0.32]$)

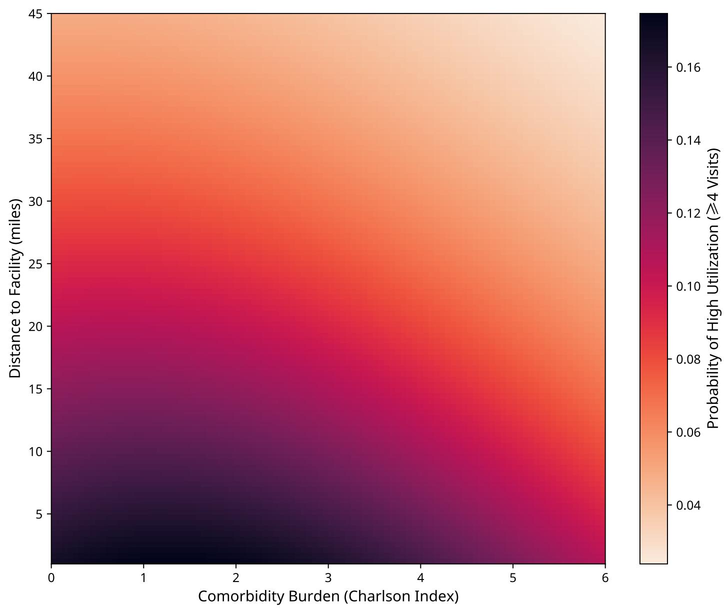

Figure 8: Healthcare Utilization Risk Profiles

suggest more predictable utilization patterns, possibly due to established care routines and comprehensive coverage. These profiles collectively demonstrate that moving beyond point estimates to interval-based predictions enables healthcare administrators to plan for realistic ranges of possible demand rather than relying on potentially misleading precise estimates, thereby supporting more resilient resource allocation, targeted intervention strategies, and ultimately more robust healthcare delivery systems that honestly acknowledge the ambiguities inherent in patient behavior.

7 Conclusion

This paper has introduced the Neutrosophic Negative Binomial Hurdle model, a novel statistical framework designed to address the critical challenge of indeterminacy in overdispersed count data with excess zeros. By integrating neutrosophic logic into the established hurdle modeling paradigm, the NNBH model explicitly quantifies uncertainty in both the binary hurdle process and the positive count mechanism, offering a more honest and flexible representation of real-world data-generating processes than classical or partially neutrosophic alternatives.

Three distinct model formulations (NNBH-I, NNBH-II, and NNBH-III) were developed to accommodate varying sources of ambiguity, whether in the hurdle probability, the count process mean, or both. The mathematical foundation of the model was rigorously established, including derivations of key statistical properties such as the neutrosophic mean, variance, hazard function, and risk measures, all of which are expressed as intervals to reflect inherent indeterminacy.

The proposed Neutrosophic Maximum Likelihood Estimation algorithm provides a practical and robust method for parameter estimation, supported by a comprehensive simulation study that confirms the consistency and asymptotic properties of the estimators. The empirical application to healthcare utilization data further validated the model's superiority, demonstrating its ability to yield more nuanced insights and better-calibrated predictions than classical or partially neutrosophic models.

From a practical standpoint, the NNBH model equips researchers and practitioners in fields such as public health, economics, and engineering with a powerful tool for analyzing complex, vague, and overdispersed count processes. By explicitly acknowledging and quantifying indeterminacy, the model supports more informed decision-making, robust risk assessment, and realistic policy planning.

Future research may explore extensions of the NNBH framework to include random effects, time-varying parameters, or Bayesian neutrosophic formulations, further enhancing its applicability to increasingly complex and dynamic data environments.

Conflict of Interest: "The author declares no conflict of interest."

References

[1] Cameron AC, Trivedi PK. *Regression analysis of count data.* 2nd ed. Cambridge: Cambridge University Press; 2013. DOI: https://doi.org/10.1017/CBO9781139013567

[2] Winkelmann R. *Econometric analysis of count data.* 5th ed. Berlin: Springer; 2008. DOI: https://doi.org/10.1007/978-3-540-78389-3

[3] Hilbe JM. *Negative binomial regression.* 2nd ed. Cambridge: Cambridge University Press; 2011. DOI: https://doi.org/10.1017/CBO9780511973420

[4] Hall DB. Zero-inflated Poisson and binomial regression with random effects: a case study. *Biometrics.* 2000; 56(4): 1030-1039. DOI: https://doi.org/10.1111/j.0006-341X.2000.01030.x

[5] Yau KK, Wang K, Lee AH. Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. *Biometrical Journal.* 2003; 45(4): 437452. DOI: https://doi.org/10.1002/bimj.200390024

[6] Omer T, Kristofer M, Sjlander P, Kibria BMG. Improved estimators for the zero-inflated Poisson regression model in the presence of multicollinearity: Simulation and applications of maternal death data. *Communications in StatisticsCase Studies and Data Analysis.* 2021; 7(3): 394412. DOI: https://doi.org/10.1080/23737484.2021.1952493

[7] Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics.* 1992; 34(1): 114. DOI: https://doi.org/10.1080/00401706.1992.10485228

[8] Greene WH. Accounting for excess zeros and sample selection in Poisson and negative binomial regression models. *Economic Working Paper EC-94-10.* New York: New York University; 1994.

[9] Nayem HM, Kibria BMG. The ridge-hurdle negative binomial regression model: A novel solution for zero-inflated counts in the presence of multicollinearity. *Stats.* 2025; 8(4): 102. DOI: https://doi.org/10.3390/stats8040102

[10] Mullahy J. Specification and testing of some modified count data models. *Journal of Econometrics.* 1986; 33(3): 341365. DOI: https://doi.org/10.1016/0304-4076(86)90002-3

[11] Arens AM, Gaher RM, Simons JS, Dvorak RD. Child maltreatment and deliberate self-harm: A negative binomial hurdle model for explanatory constructs. *Child Maltreatment.* 2014; 19(34): 168177. DOI: https://doi.org/10.1177/1077559514548315

[12] Smarandache F. Neutrosophic seta generalization of the intuitionistic fuzzy set. *International Journal of Pure and Applied Mathematics.* 2005; 24(3): 287297.

[13] Smarandache F. *Introduction to neutrosophic statistics.* Craiova: Sitech and Education Publisher; Columbus, OH: Educational Publisher; 2014. DOI: https://doi.org/10.13140/2.1.2780.1289

[14] Smarandache F. *A unifying field in logics: neutrosophic logic, neutrosophic set, neutrosophic probability.* Rehoboth: American Research Press; 1999.

[15] Rao GS. Neutrosophic log-logistic distribution model in complex alloy metal melting point applications. *International Journal of Computational Intelligence Systems.* 2023; 16(1): 48. DOI: https://doi.org/10.1007/s44196-023-00218-5

[16] Lathamaheswari M, Broumi S, Smarandache F, Sudha S. Neutrosophic perspective of neutrosophic probability distributions and its application. *International Journal of Neutrosophic Science.* 2021; 17(2): 96-109. DOI: https://doi.org/10.54216/IJNS.170202

[17] Alhabib R, Ranna MM, Farah H, Salama AA. Some neutrosophic probability distributions. *Neutrosophic Sets and Systems.* 2018; 22: 30-38. DOI: https://doi.org/10.5281/zenodo.2160478

[18] Aslam M. The neutrosophic negative binomial distribution: algorithms and practical application. *REVSTATStatistical Journal.* 2024.

[19] Alanzi AR, Irshad MR, Johny M, Al-Omari AI, Alrweili H. Neutrosophic Poisson moment exponential distribution: Properties and applications. *Maejo International Journal of Science and Technology.* 2025; 19(02): 133-149. DOI: https://doi.org/10.14456/mijst.2025.11

[20] Al-Essa LA, Jamal F, Shafiq S, Khan S, Abbas Q, Khan Sherwani RA, Aslam M. Properties and applications of neutrosophic Burr XII distribution. *International Journal of Computational Intelligence Systems.* 2025; 18(1): 10. DOI: https://doi.org/10.1007/s44196-024-00721-3

[21] Sadok I, Zribi M, Masmoudi A. Non-informative Bayesian estimation in dispersion models. *Hacettepe Journal of Mathematics and Statistics.* 2023; 53(1): 251-268. DOI: https://doi.org/10.15672/hujms.1053432

[22] Sadok I, Masmoudi A, Zribi M. Integrating the EM algorithm with particle filter for image restoration with exponential dispersion noise. *Communications in StatisticsTheory and Methods.* 2023; 52(2): 446-462. DOI: https://doi.org/10.1080/03610926.2021.1915336

Ibrahim Sadok
Department of Mathematics and Computer Science
University of Bechar, Faculty of Exact Sciences
Bechar, Algeria
E-mail: ibrahim.sadok@univ-bechar.dz