Statistical Inference for Series Systems from Masked Failure Time Data: The Exponential Case

Alexander Towell

`atowell@siue.edu`

Abstract

We consider the problem of estimating component failure rates in series systems when observations consist of system failure times paired with partial information about the failed component. For the case where component lifetimes follow exponential distributions, we derive closed-form expressions for the maximum likelihood estimator, the Fisher information matrix, and establish sufficient statistics. The asymptotic sampling distribution of the estimator is characterized and confidence intervals are provided. A detailed analysis of a three-component system demonstrates the theoretical results.

Contents

1   Introduction

Series systems are fundamental in reliability engineering: the system fails whenever any single component fails. In many practical situations, the exact cause of system failure cannot be determined with certainty, but can be narrowed to a subset of components. This partial information is known as *masked* failure data. For example, in electronic systems a diagnostic test might isolate a failure to one of two circuit boards without determining which board actually failed; in medical devices, system failure might be attributed to a subset of components based on failure mode symptoms; or in aerospace

systems, post-failure analysis might narrow the cause to a specific subsystem but cannot identify the individual component responsible.

The problem arises in two common scenarios: (1) diagnostic limitations where identifying the exact failed component requires destructive testing or prohibitive expense, and (2) field data collection where only partial diagnostic information is recorded. Despite this incomplete information, engineers must estimate component-level reliability parameters to make design improvements, optimize maintenance schedules, and predict system lifetime distributions.

1.1 Related Work

The analysis of masked failure data has its roots in competing risks theory. Cox [4] introduced the latent failure time model for analyzing exponentially distributed lifetimes with multiple failure types, establishing the foundational framework for modeling situations where multiple competing causes may lead to system failure. This work demonstrated that under exponential assumptions, the failed component identity and system lifetime can be independent random variables, a property that proves crucial for the analysis in this paper.

Miyakawa [12] pioneered the application of competing risks methods specifically to masked system failure data, deriving closed-form maximum likelihood estimates for a two-component series system with exponential component lifetimes. Usher and Hodgson [15] extended this approach, introducing general maximum likelihood methods for estimating component reliability from masked system life-test data and providing computational procedures for series systems. Lin et al. [11] further developed exact maximum likelihood estimation procedures for three-component systems under more general masking scenarios.

Dinse [5] developed nonparametric methods for partially-complete time and type of failure data, introducing an iterative algorithm yielding distribution-free estimates that converge to maximum likelihood solutions. This work established that useful inference is possible even when the failure cause is only partially observed.

The 1990s saw significant development of Bayesian approaches to masked data. Reiser et al. [13] introduced Bayesian inference methods for masked system lifetime data, while Guttman et al. [9] addressed dependent masking scenarios where the probability of masking depends on the true cause of failure, developing Bayesian methodology for two-component systems. These

Bayesian approaches provided posterior distributions for component reliabilities but did not derive closed-form expressions for frequentist estimators or their asymptotic properties.

Flehinger et al. [6, 7] contributed to the theoretical foundations by studying survival analysis with competing risks and masked failure causes, developing parametric models that accommodate general patterns of missing failure types. Sarhan [14] provided both maximum likelihood and Bayes procedures for estimating component reliabilities in series systems with exponentially distributed lifetimes, focusing on practical implementation issues.

More recent work has expanded the scope to include interval-censored data [8], Bayesian modeling frameworks [10], and applications to specific engineering contexts. The competing risks framework continues to provide a natural setting for masked data analysis [1], connecting reliability engineering with survival analysis and biostatistics.

Despite this substantial body of work, *closed-form analytical results remain limited*. Specifically:

- The Fisher information matrix for masked system data has not been derived in closed form for general series systems, even under exponential assumptions

- Asymptotic properties of maximum likelihood estimators have not been fully characterized without numerical computation

- Minimal sufficient statistics have not been identified for general masking patterns

- Most existing methods rely on numerical optimization or EM algorithms without explicit characterization of estimator variance and covariance structure

This paper addresses these gaps by providing complete analytical results for the exponential case. The exponential distribution is a cornerstone of reliability theory [2], justified in several contexts: systems subject to random external shocks, systems with constant hazard rates, and early-life failure analysis where wear-out has not yet begun. The exponential assumption also serves as a limiting case for other distributions and provides a baseline for comparison with more complex models. More importantly, the exponential case admits closed-form solutions that provide direct insight into the information structure of masked failure data.

1.2 Contributions

This paper provides the first complete analytical treatment of maximum likelihood estimation for series systems with masked failure data under exponential component lifetimes. Our main contributions are:

1. **Explicit Fisher information matrix**: We derive a closed-form expression for the Fisher information matrix under arbitrary masking patterns, enabling direct computation of asymptotic variances without numerical differentiation or Monte Carlo simulation

2. **Sufficient statistics**: The mean system lifetime and candidate set frequency vector constitute sufficient statistics, reducing the data to $1 + \binom{m}{w}$ real numbers

3. **Closed-form MLE for $w = m-1$ masking**: For masking cardinality $w = m-1$ (where each candidate set excludes exactly one component), we derive an explicit closed-form solution to the likelihood equations for arbitrary m, eliminating the need for numerical optimization. This represents the first known closed-form MLE for non-trivial masking scenarios. The three-component case ($m = 3$, $w = 2$) is developed in detail

4. **Asymptotic distribution theory**: We characterize the asymptotic sampling distribution of the MLE and provide Wald-type confidence intervals with explicit formulas for coverage

5. **Numerical validation**: We provide Monte Carlo evidence confirming the theoretical asymptotic covariance matches empirical behavior for finite samples

The closed-form results for $w = m-1$ are particularly significant because this configuration represents the simplest non-degenerate masking scenario: when $w = 1$ there is no masking (exact cause identification), and when $w = m$ there is complete masking (no diagnostic information). The $w = m-1$ case captures the essential statistical structure of masked data while admitting analytical solutions.

1.3 Paper organization

Section 2 introduces the probabilistic model for series systems with masked failures. Section 3 develops the likelihood and Fisher information for general parametric families. Section 4 presents the main results for exponentially distributed components, including the MLE, information matrix, and sufficient statistics. Section 5 provides detailed analysis of three-component systems. Section 6 concludes with discussion.

2 Probabilistic Model

2.1 Series System Lifetime

Consider a system composed of m components. Component j has a random lifetime $T_j > 0$ for $j = 1, \ldots, m$. We make the following assumptions:

Assumption 2.1. The component lifetimes $T_1, \ldots, T_m$ are mutually independent.

Assumption 2.2. The system operates if and only if all components are functioning (series configuration).

Under these assumptions, the system lifetime is:

$$S = \min(T_1, \ldots, T_m) \tag{1}$$

Let $F_j(t)$ and $f_j(t)$ denote the CDF and PDF of component j. Define the reliability function $R_j(t) = 1 - F_j(t) = P\{T_j > t\}$.

The system reliability function is:

$$R_S(t) = \prod_{j=1}^{m} R_j(t) \tag{2}$$

The system PDF is:

$$f_S(t) = \sum_{j=1}^{m} \left(f_j(t) \prod_{\substack{k=1 \\ k \neq j}}^{m} R_k(t) \right) \tag{3}$$

6

2.2 Masked Component Failures

When the system fails at time t, exactly one component caused the failure. Let $K \in \{1, \ldots, m\}$ denote the failed component. In many applications, K cannot be observed directly. Instead, we observe a *candidate set* $C \subseteq \{1, \ldots, m\}$ that contains the failed component but does not uniquely identify it.

Definition 2.3. Under *uniform masking* with cardinality w, given that component k failed, the observed candidate set C is a uniformly random subset of size w containing k. That is,

$$P\{C = c | K = k\} = \begin{cases} \frac{1}{\binom{m-1}{w-1}} & \text{if } k \in c \text{ and } |c| = w \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Remark 2.4 (Interpretation and Implications). The uniform masking model has two key components:

1. The failed component k is *always* included in the candidate set ($P\{k \in C | K = k\} = 1$)

2. Each non-failed component $j \neq k$ is included in C with *equal probability* $\frac{w-1}{m-1}$, independently

This is a **strong assumption** that is generally *not satisfied in practice*. In real diagnostic scenarios:

- Components with similar failure modes may be systematically grouped together

- Accessible components may be more likely to be identified than inaccessible ones

- Expensive tests may create diagnostic biases

- Prior failure history may influence which components are suspected

Remark 2.5 (Justification for Uniform Masking). We adopt the uniform masking assumption because it enables *analytical derivation* of closed-form expressions for the MLE, Fisher information matrix, and sufficient statistics. These analytical results provide:

- Direct insight into the information structure of masked failure data

- Closed-form asymptotic variance formulas (no Monte Carlo needed)

- Baseline performance bounds for comparison with more complex masking models

- Computational efficiency (no numerical optimization needed for $m = 3$, $w = 2$)

The uniform masking model is most appropriate when:

- The diagnostic process randomly samples components for inspection

- The masking mechanism is designed to be unbiased (e.g., automated random selection)

- No prior information about component reliabilities is used during diagnosis

Despite its limitations, the uniform masking model serves as a tractable *reference model* for understanding masked failure data, analogous to how the exponential distribution serves as a reference for lifetime modeling despite the restrictive constant-hazard assumption.

Assumption 2.6. The masking mechanism (which components are included in C) is independent of the system lifetime S and depends only on the failed component K.

This assumption implies that the masking quality does not depend on when the failure occurred, only on which component failed and the diagnostic capabilities of the inspection process.

Definition 2.7. A *masked system failure time* is a pair (t, C) where t is the observed system failure time and C is the candidate set. When the masking cardinality $w = |C|$ is fixed by experimental design, we condition on w throughout.

2.3 Parametric Families

We assume component j has a lifetime distribution from a parametric family indexed by parameter $\theta_j^\star$. The system parameter is $\Theta^\star = (\theta_1^\star, \ldots, \theta_m^\star)$.

Our data consists of a random sample of n independent masked system failure times:

$$\mathbf{M}_n = \{(t_1, C_1), \ldots, (t_n, C_n)\} \tag{5}$$

where we condition on fixed cardinality w for simplicity.

3 Likelihood and Fisher Information

3.1 Likelihood Function

Given the model assumptions, the conditional probability that component k failed given system failure at time t is:

$$p_{K|S}(k|t, \Theta^\star) = \frac{f_k(t) \prod_{j \neq k} R_j(t)}{f_S(t)} \tag{6}$$

Under the uniform masking model (Definition 2.3) with fixed cardinality w, we derive the conditional probability of observing candidate set C given system failure at time t.

Proposition 3.1. *Under uniform masking with cardinality w, the conditional probability of observing candidate set C given system failure at time t is:*

$$p_{C|S,W}(C|t, w, \Theta^\star) = \frac{\sum_{j \in C} f_j(t) \prod_{k \neq j} R_k(t)}{\binom{m-1}{w-1} f_S(t)} \tag{7}$$

Proof. By the law of total probability and independence of masking from system lifetime (Assumption 2.6):

$$
\begin{aligned}
p_{C|S}(C|t, w) &= \sum_{k=1}^{m} p_{C|K,S}(C|k, t, w) \cdot p_{K|S}(k|t) \\
&= \sum_{k=1}^{m} p_{C|K}(C|k, w) \cdot p_{K|S}(k|t) \tag{8}
\end{aligned}
$$

Under uniform masking (Definition 2.3):

$$p_{C|K}(C|k, w) = \begin{cases} \frac{1}{\binom{m-1}{w-1}} & \text{if } k \in C \\ 0 & \text{if } k \notin C \end{cases} \tag{9}$$

Therefore:

$$p_{C|S}(C|t, w) = \sum_{k \in C} \frac{1}{\binom{m-1}{w-1}} \cdot \frac{f_k(t) \prod_{j \neq k} R_j(t)}{f_S(t)}$$

$$= \frac{1}{\binom{m-1}{w-1}} \cdot \frac{\sum_{k \in C} f_k(t) \prod_{j \neq k} R_j(t)}{f_S(t)} \tag{10}$$

$\square$

This result shows that the candidate set probability is proportional to the sum of individual component failure probabilities over all components in the candidate set, normalized by the binomial coefficient representing the number of candidate sets containing each component.

The joint density of system failure time and candidate set is:

$$f_{C,S|W}(C, t|w, \Theta^\star) = \frac{1}{\binom{m-1}{w-1}} \sum_{j \in C} f_j(t) \prod_{k \neq j} R_k(t) \tag{11}$$

The likelihood function for sample $\mathbf{M}_n$ is:

$$\mathcal{L}(\Theta|\mathbf{M}_n) = \prod_{i=1}^{n} f_{C,S|W}(C_i, t_i|w, \Theta) \tag{12}$$

3.2 Fisher Information Matrix

The Fisher information matrix quantifies the expected information about $\Theta^\star$ contained in a single observation. Under regularity conditions, the (i, j)-th element is:

$$[\mathcal{I}(\Theta^\star|w)]_{ij} = -E\{\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln f_{C,S|W}(C, S|w, \Theta)\}[\Theta^\star] \tag{13}$$

For a sample of n observations, the information is additive:

$$\mathcal{I}_n(\Theta^\star|w) = n \cdot \mathcal{I}(\Theta^\star|w) \tag{14}$$

10

4 Exponentially Distributed Component Lifetimes

4.1 Exponential Parametric Functions

Suppose component j has an exponentially distributed lifetime with failure rate $\lambda_j^\star$, denoted $\mathrm{T}_j \sim \mathrm{EXP}(\lambda_j^\star)$. The component has:

$$R_j(t|\lambda_j^\star) = \exp(-\lambda_j^\star t) \tag{15}$$
$$f_j(t|\lambda_j^\star) = \lambda_j^\star \exp(-\lambda_j^\star t) \tag{16}$$
$$h_j(t|\lambda_j^\star) = \lambda_j^\star \tag{17}$$

where $t > 0$ and $\lambda_j^\star > 0$.

Theorem 4.1. *A series system with exponentially distributed component lifetimes is exponentially distributed with failure rate $\sum_{j=1}^m \lambda_j^\star$.*

Proof. By equation (2),

$$R_\mathrm{S}(t|\boldsymbol{\lambda}^\star) = \prod_{j=1}^m \exp(-\lambda_j^\star t) = \exp\left(-\left[\sum_{j=1}^m \lambda_j^\star\right] t\right) \tag{18}$$

which is the reliability function of an exponential distribution with rate $\sum_{j=1}^m \lambda_j^\star$. $\qquad\square$

The system PDF is:

$$f_\mathrm{S}(t|\boldsymbol{\lambda}^\star) = \left(\sum_{j=1}^m \lambda_j^\star\right) \exp\left(-\left[\sum_{j=1}^m \lambda_j^\star\right] t\right) \tag{19}$$

Proposition 4.2. *For exponential series systems, the failed component* K *and system lifetime* S *are independent. The marginal distribution of* K *is:*

$$p_\mathrm{K}(k|\boldsymbol{\lambda}^\star) = \frac{\lambda_k^\star}{\sum_{j=1}^m \lambda_j^\star} \tag{20}$$

Proof. Let $\Lambda = \sum_{j=1}^m \lambda_j^\star$. We compute:

$$\mathrm{P}\{\mathrm{K} = k, \mathrm{S} \leq t\} = \int_0^t \lambda_k^\star \exp(-\Lambda s)\,ds = \frac{\lambda_k^\star}{\Lambda}\left(1 - e^{-\Lambda t}\right)$$
$$= \mathrm{P}\{\mathrm{K} = k\} \cdot \mathrm{P}\{\mathrm{S} \leq t\} \tag{21}$$

since $\mathrm{P}\{\mathrm{K} = k\} = \lambda_k^\star/\Lambda$ and $\mathrm{P}\{\mathrm{S} \leq t\} = 1 - e^{-\Lambda t}$. $\qquad\square$

Remark 4.3. Intuitively, the memoryless property of exponential distributions means that at any instant, each component has the same probability of being the next to fail, proportional to its rate λ_j. The *when* (system lifetime) and *which* (failed component) are determined by independent processes. This independence is specific to exponential distributions; for Weibull or other distributions with time-varying hazard rates, K and S are generally dependent.

The joint density is:

$$f_{K,S}(k, t|\boldsymbol{\lambda}^\star) = \lambda_k^\star \exp\left(-\left[\sum_{j=1}^m \lambda_j^\star\right] t\right) \tag{22}$$

Under the uniform masking model (Definition 2.3), the joint density of candidate set and system failure time is:

$$f_{C,S|W}(C, t|w, \boldsymbol{\lambda}^\star) = \frac{1}{\binom{m-1}{w-1}} \left(\sum_{j \in C} \lambda_j^\star\right) \exp\left(-\left[\sum_{j=1}^m \lambda_j^\star\right] t\right) \tag{23}$$

This follows directly from Proposition 3.1 by substituting the exponential PDFs and noting that for exponential distributions, $\prod_{k \neq j} R_k(t) = \exp(-\sum_{k \neq j} \lambda_k^\star t)$.

4.2 Maximum Likelihood Estimator

The likelihood function for sample $\mathbf{M}_n$ with candidate sets of cardinality w is:

$$\mathcal{L}(\boldsymbol{\lambda}|\mathbf{M}_n) \propto \exp\left(-\left[\sum_{j=1}^m \lambda_j\right]\left[\sum_{i=1}^n t_i\right]\right) \prod_{i=1}^n \left(\sum_{j \in C_i} \lambda_j\right) \tag{24}$$

The log-likelihood is:

$$\ell(\boldsymbol{\lambda}|\mathbf{M}_n) = \sum_{i=1}^n \ln\left(\sum_{j \in C_i} \lambda_j\right) - \left[\sum_{j=1}^m \lambda_j\right]\left[\sum_{i=1}^n t_i\right] \tag{25}$$

The score function has j-th component:

$$\frac{\partial \ell}{\partial \lambda_j} = \sum_{i=1}^n \frac{\mathbb{1}_{C_i}(j)}{\sum_{k \in C_i} \lambda_k} - \sum_{i=1}^n t_i \tag{26}$$

where $\mathbb{1}_C(j)$ is the indicator function equal to 1 if $j \in C$ and 0 otherwise.

Theorem 4.4. *The maximum likelihood estimator $\hat{\boldsymbol{\lambda}}_n$ maximizes the log-likelihood* (25) *and satisfies the score equation $\nabla\ell(\hat{\boldsymbol{\lambda}}_n|\mathbf{M}_n) = \mathbf{0}$.*

In general, this requires numerical solution. However, for specific cases (notably three-component systems with $w = 2$), closed-form solutions exist.

4.3 Sufficient Statistics

Theorem 4.5. *For masked system failure times from exponentially distributed series systems, the statistics*

$$\bar{t} = \frac{1}{n}\sum_{i=1}^{n} t_i \quad and \quad \hat{\boldsymbol{\omega}} = (\hat{\boldsymbol{\omega}}_{\mathrm{C}})_{\mathrm{C}} \tag{27}$$

where $\hat{\boldsymbol{\omega}}_{\mathrm{C}}$ denotes the count of observations with candidate set C, for each of the $\binom{m}{w}$ possible candidate sets, are jointly sufficient for $\boldsymbol{\lambda}^{\star}$.

Proof. The likelihood can be factored as:

$$\mathcal{L}(\boldsymbol{\lambda}|\mathbf{M}_n) = \exp\left(-n\bar{t}\sum_{j=1}^{m}\lambda_j\right)\prod_{\mathrm{C}}\left(\sum_{j\in\mathrm{C}}\lambda_j\right)^{\hat{\boldsymbol{\omega}}_{\mathrm{C}}} \tag{28}$$

which depends on the data only through $\bar{t}$ and $\hat{\boldsymbol{\omega}}$. By the factorization theorem, these are sufficient statistics. $\square$

The sufficiency result shows that all information about $\boldsymbol{\lambda}^{\star}$ in the sample is captured by: (1) the average system lifetime, and (2) the frequencies of each candidate set. We do not establish minimality: the likelihood is not a regular exponential family in $\boldsymbol{\lambda}$ because the natural parameters involve $\ln(\sum_{j\in\mathrm{C}}\lambda_j)$, so a separate argument would be needed to show that no further reduction is possible.

4.4 Fisher Information Matrix

The (j, k)-th element of the Fisher information matrix for exponential series systems is:

$$[\mathcal{I}(\boldsymbol{\lambda}^{\star}|w)]_{jk} = \frac{\sum_{\mathrm{C}}\left(\sum_{p\in\mathrm{C}}\lambda_p^{\star}\right)^{-1}\mathbb{1}_{\mathrm{C}\times\mathrm{C}}(j, k)}{\binom{m-1}{w-1}\sum_{p=1}^{m}\lambda_p^{\star}} \tag{29}$$

13

where the sum is over all candidate sets C of cardinality w.

This formula reveals that information about the pair (λ_j, λ_k) accrues only from candidate sets containing both components j and k (the indicator $\mathbb{1}_{\mathrm{C} \times \mathrm{C}}(j, k)$ ensures this). Each such candidate set contributes inversely to its total failure rate, so components with higher failure rates contribute less information per observation. The denominator scales with the total failure rate and the number of possible candidate sets.

Proof. From equation (23), the log-density is:

$$\ln f_{\mathrm{C,S|W}}(\mathrm{C}, t | w, \boldsymbol{\lambda}) = \ln \left(\sum_{p \in \mathrm{C}} \lambda_p \right) - \left(\sum_{q=1}^{m} \lambda_q \right) t + \text{const} \qquad (30)$$

Taking the first partial derivative with respect to λ_j:

$$\frac{\partial}{\partial \lambda_j} \ln f_{\mathrm{C,S|W}} = \frac{\mathbb{1}_{\mathrm{C}}(j)}{\sum_{p \in \mathrm{C}} \lambda_p} - t \qquad (31)$$

where $\mathbb{1}_{\mathrm{C}}(j) = 1$ if $j \in \mathrm{C}$ and 0 otherwise.

Taking the second partial derivative with respect to λ_k:

$$\begin{aligned}
\frac{\partial^2}{\partial \lambda_j \partial \lambda_k} \ln f_{\mathrm{C,S|W}} &= \frac{\partial}{\partial \lambda_k} \left[\frac{\mathbb{1}_{\mathrm{C}}(j)}{\sum_{p \in \mathrm{C}} \lambda_p} \right] \\
&= -\frac{\mathbb{1}_{\mathrm{C}}(j) \cdot \mathbb{1}_{\mathrm{C}}(k)}{(\sum_{p \in \mathrm{C}} \lambda_p)^2} \\
&= -\frac{\mathbb{1}_{\mathrm{C} \times \mathrm{C}}(j, k)}{(\sum_{p \in \mathrm{C}} \lambda_p)^2} \qquad (32)
\end{aligned}$$

By definition of Fisher information:

$$\begin{aligned}
[\mathcal{I}(\boldsymbol{\lambda}^\star | w)]_{jk} &= -\mathrm{E}\{\frac{\partial^2}{\partial \lambda_j \partial \lambda_k} \ln f_{\mathrm{C,S|W}}\}[\boldsymbol{\lambda}^\star] \\
&= \sum_{\mathrm{C}:|\mathrm{C}|=w} \int_0^\infty \frac{\mathbb{1}_{\mathrm{C} \times \mathrm{C}}(j, k)}{(\sum_{p \in \mathrm{C}} \lambda_p^\star)^2} \cdot f_{\mathrm{C,S|W}}(\mathrm{C}, t | w, \boldsymbol{\lambda}^\star) \, dt \qquad (33)
\end{aligned}$$

14

Substituting the joint density from equation (23):

$$[\mathcal{I}(\boldsymbol{\lambda}^\star|w)]_{jk} = \sum_{\mathrm{C}:|\mathrm{C}|=w} \frac{\mathbb{1}_{\mathrm{C}\times\mathrm{C}}(j,k)}{(\sum_{p\in\mathrm{C}}\lambda_p^\star)^2} \cdot \frac{1}{\binom{m-1}{w-1}}$$
$$\times \int_0^\infty \left(\sum_{p\in\mathrm{C}}\lambda_p^\star\right) \exp\left(-\left[\sum_{q=1}^m \lambda_q^\star\right]t\right) dt \qquad (34)$$

Evaluating the integral using $\int_0^\infty e^{-at}dt = 1/a$ for $a > 0$:

$$\int_0^\infty \left(\sum_{p\in\mathrm{C}}\lambda_p^\star\right) \exp\left(-\left[\sum_{q=1}^m \lambda_q^\star\right]t\right) dt = \frac{\sum_{p\in\mathrm{C}}\lambda_p^\star}{\sum_{q=1}^m \lambda_q^\star} \qquad (35)$$

Therefore:

$$[\mathcal{I}(\boldsymbol{\lambda}^\star|w)]_{jk} = \sum_{\mathrm{C}:|\mathrm{C}|=w} \frac{\mathbb{1}_{\mathrm{C}\times\mathrm{C}}(j,k)}{(\sum_{p\in\mathrm{C}}\lambda_p^\star)^2} \cdot \frac{1}{\binom{m-1}{w-1}} \cdot \frac{\sum_{p\in\mathrm{C}}\lambda_p^\star}{\sum_{q=1}^m \lambda_q^\star}$$
$$= \frac{1}{\binom{m-1}{w-1}\sum_{q=1}^m \lambda_q^\star} \sum_{\mathrm{C}:|\mathrm{C}|=w} \frac{\mathbb{1}_{\mathrm{C}\times\mathrm{C}}(j,k)}{\sum_{p\in\mathrm{C}}\lambda_p^\star} \qquad (36)$$

This completes the derivation of equation (29). $\qquad\qquad\square$

Remark 4.6 (Combining Estimates from Variable Masking Cardinality). In practice, diagnostic quality may vary across observations, resulting in data with different masking cardinalities. Suppose the sample is partitioned into G groups, where group g has n_g observations with masking cardinality w_g. Each group yields an MLE $\hat{\boldsymbol{\lambda}}^{(g)}$ with asymptotic covariance $\frac{1}{n_g}\mathcal{I}^{-1}(\boldsymbol{\lambda}^\star|w_g)$.

The optimal combined estimator under independence is the inverse-variance weighted average:

$$\hat{\boldsymbol{\lambda}}_{\mathrm{combined}} = \left(\sum_{g=1}^G n_g\mathcal{I}(\boldsymbol{\lambda}^\star|w_g)\right)^{-1} \left(\sum_{g=1}^G n_g\mathcal{I}(\boldsymbol{\lambda}^\star|w_g)\hat{\boldsymbol{\lambda}}^{(g)}\right) \qquad (37)$$

with asymptotic covariance

$$\mathrm{Cov}(\hat{\boldsymbol{\lambda}}_{\mathrm{combined}}) = \left(\sum_{g=1}^G n_g\mathcal{I}(\boldsymbol{\lambda}^\star|w_g)\right)^{-1} \qquad (38)$$

15

This combined estimator achieves minimum variance among all linear unbiased combinations. In practice, the true parameter $\boldsymbol{\lambda}^\star$ in the information matrices is replaced by the combined estimate $\hat{\boldsymbol{\lambda}}_{\text{combined}}$, yielding a feasible estimator.

4.5 Asymptotic Sampling Distribution

The following regularity conditions ensure consistency and asymptotic normality of the MLE:

(R1) The true parameter $\boldsymbol{\lambda}^\star$ lies in the interior of the parameter space $\Theta = (0, \infty)^m$

(R2) The Fisher information matrix $\mathcal{I}(\boldsymbol{\lambda}^\star|w)$ is positive definite

(R3) The observations $(t_i, C_i)_{i=1}^n$ are independent and identically distributed

Condition (R1) excludes boundary cases where components have zero failure rate. Condition (R2) is satisfied when each component appears in at least one candidate set, ensuring all parameters are identifiable. Condition (R3) holds by the experimental design.

Theorem 4.7. *Under conditions (R1)–(R3), as $n \to \infty$,*

$$\sqrt{n}(\hat{\boldsymbol{\lambda}}_n - \boldsymbol{\lambda}^\star) \xrightarrow{d} \text{MVN}\left(\mathbf{0}, \mathcal{I}^{-1}(\boldsymbol{\lambda}^\star|w)\right) \tag{39}$$

This result follows from standard maximum likelihood theory [3]. The asymptotic variance-covariance matrix is the inverse of the Fisher information matrix.

4.6 Confidence Intervals

An asymptotic $(1 - \alpha) \times 100\%$ confidence interval for $\lambda_j^\star$ is:

$$\hat{\lambda}_j \pm z_{1-\alpha/2}\sqrt{\frac{1}{n}[\mathcal{I}^{-1}(\hat{\boldsymbol{\lambda}}_n|w)]_{jj}} \tag{40}$$

where $z_{1-\alpha/2}$ is the $(1 - \alpha/2)$-quantile of the standard normal distribution.

5 Three-Component Systems

We provide detailed analysis for systems with $m = 3$ components, which admits several closed-form results.

5.1 Candidate Sets of Size Two

Consider observations where each candidate set has cardinality $w = m - 1$, meaning exactly one component is excluded from each candidate set. For general m, a closed-form MLE exists.

Theorem 5.1. *For m-component systems with masking cardinality $w = m - 1$, define for each component j:*

- $A_j = \sum_{C \ni j} \hat{\boldsymbol{\omega}}_C$: *the sum of candidate set counts containing component j*

- $B_j = \hat{\boldsymbol{\omega}}_{\{1,\ldots,m\}\setminus\{j\}}$: *the count for the unique candidate set excluding component j*

The MLE has the closed-form solution:

$$\hat{\lambda}_j = \frac{A_j - (m-2)B_j}{n\bar{t}} \tag{41}$$

For three-component systems ($m = 3$) with $w = 2$, the candidate sets are $\{1, 2\}$, $\{1, 3\}$, and $\{2, 3\}$.

Corollary 5.2. *For three-component systems with $w = 2$, the MLE has the closed-form solution:*

$$\hat{\boldsymbol{\lambda}}_n = \frac{1}{n\bar{t}} \begin{pmatrix} \hat{\boldsymbol{\omega}}_{\{1,2\}} + \hat{\boldsymbol{\omega}}_{\{1,3\}} - \hat{\boldsymbol{\omega}}_{\{2,3\}} \\ \hat{\boldsymbol{\omega}}_{\{1,2\}} - \hat{\boldsymbol{\omega}}_{\{1,3\}} + \hat{\boldsymbol{\omega}}_{\{2,3\}} \\ -\hat{\boldsymbol{\omega}}_{\{1,2\}} + \hat{\boldsymbol{\omega}}_{\{1,3\}} + \hat{\boldsymbol{\omega}}_{\{2,3\}} \end{pmatrix}. \tag{42}$$

Proof of Corollary 5.2. For $m = 3$: $A_1 = \hat{\boldsymbol{\omega}}_{\{1,2\}} + \hat{\boldsymbol{\omega}}_{\{1,3\}}$, $B_1 = \hat{\boldsymbol{\omega}}_{\{2,3\}}$, and $(m-2) = 1$. Applying Theorem 5.1: $\hat{\lambda}_1 = (A_1 - B_1)/(n\bar{t}) = (\hat{\boldsymbol{\omega}}_{\{1,2\}} + \hat{\boldsymbol{\omega}}_{\{1,3\}} - \hat{\boldsymbol{\omega}}_{\{2,3\}})/(n\bar{t})$. The formulas for $\hat{\lambda}_2$ and $\hat{\lambda}_3$ follow symmetrically. $\square$

The log-likelihood for the three-component case is:

$$\ell(\boldsymbol{\lambda}|\bar{t},\hat{\boldsymbol{\omega}}) = \hat{\boldsymbol{\omega}}_{\{1,2\}}\ln(\lambda_1+\lambda_2)+\hat{\boldsymbol{\omega}}_{\{1,3\}}\ln(\lambda_1+\lambda_3)+\hat{\boldsymbol{\omega}}_{\{2,3\}}\ln(\lambda_2+\lambda_3)-n\bar{t}(\lambda_1+\lambda_2+\lambda_3)$$
(43)

The score equations are:

$$\nabla\ell = \begin{pmatrix} \frac{\hat{\boldsymbol{\omega}}_{\{1,2\}}}{\lambda_1+\lambda_2} + \frac{\hat{\boldsymbol{\omega}}_{\{1,3\}}}{\lambda_1+\lambda_3} \\ \frac{\hat{\boldsymbol{\omega}}_{\{1,2\}}}{\lambda_1+\lambda_2} + \frac{\hat{\boldsymbol{\omega}}_{\{2,3\}}}{\lambda_2+\lambda_3} \\ \frac{\hat{\boldsymbol{\omega}}_{\{1,3\}}}{\lambda_1+\lambda_3} + \frac{\hat{\boldsymbol{\omega}}_{\{2,3\}}}{\lambda_2+\lambda_3} \end{pmatrix} - n\bar{t}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \mathbf{0}$$
(44)

Proof of Theorem 5.1 for $m = 3$. We demonstrate the result for the three-component case; the extension to arbitrary m follows by analogous algebra. The key insight is that pairwise subtraction of score equations yields proportionality relations among candidate set sums.

Setting the score equations to zero:

$$\frac{\hat{\boldsymbol{\omega}}_{\{1,2\}}}{\lambda_1 + \lambda_2} + \frac{\hat{\boldsymbol{\omega}}_{\{1,3\}}}{\lambda_1 + \lambda_3} = n\bar{t}$$
(45)

$$\frac{\hat{\boldsymbol{\omega}}_{\{1,2\}}}{\lambda_1 + \lambda_2} + \frac{\hat{\boldsymbol{\omega}}_{\{2,3\}}}{\lambda_2 + \lambda_3} = n\bar{t}$$
(46)

$$\frac{\hat{\boldsymbol{\omega}}_{\{1,3\}}}{\lambda_1 + \lambda_3} + \frac{\hat{\boldsymbol{\omega}}_{\{2,3\}}}{\lambda_2 + \lambda_3} = n\bar{t}$$
(47)

Define the pairwise sums: $\sigma_1 = \lambda_1 + \lambda_2$, $\sigma_2 = \lambda_1 + \lambda_3$, $\sigma_3 = \lambda_2 + \lambda_3$. Then the score equations become:

$$\frac{\hat{\boldsymbol{\omega}}_{\{1,2\}}}{\sigma_1} + \frac{\hat{\boldsymbol{\omega}}_{\{1,3\}}}{\sigma_2} = n\bar{t}$$
(48)

$$\frac{\hat{\boldsymbol{\omega}}_{\{1,2\}}}{\sigma_1} + \frac{\hat{\boldsymbol{\omega}}_{\{2,3\}}}{\sigma_3} = n\bar{t}$$
(49)

$$\frac{\hat{\boldsymbol{\omega}}_{\{1,3\}}}{\sigma_2} + \frac{\hat{\boldsymbol{\omega}}_{\{2,3\}}}{\sigma_3} = n\bar{t}$$
(50)

Subtracting equation (49) from (48):

$$\frac{\hat{\boldsymbol{\omega}}_{\{1,3\}}}{\sigma_2} - \frac{\hat{\boldsymbol{\omega}}_{\{2,3\}}}{\sigma_3} = 0 \quad \Rightarrow \quad \hat{\boldsymbol{\omega}}_{\{1,3\}}\sigma_3 = \hat{\boldsymbol{\omega}}_{\{2,3\}}\sigma_2$$
(51)

Subtracting equation (50) from (48):

$$\frac{\hat{\boldsymbol{\omega}}_{\{1,2\}}}{\sigma_1} - \frac{\hat{\boldsymbol{\omega}}_{\{2,3\}}}{\sigma_3} = 0 \quad \Rightarrow \quad \hat{\boldsymbol{\omega}}_{\{1,2\}}\sigma_3 = \hat{\boldsymbol{\omega}}_{\{2,3\}}\sigma_1$$
(52)

18

Subtracting equation (50) from (49):

$$\frac{\hat{\boldsymbol{\omega}}_{\{1,2\}}}{\sigma_1} - \frac{\hat{\boldsymbol{\omega}}_{\{1,3\}}}{\sigma_2} = 0 \quad \Rightarrow \quad \hat{\boldsymbol{\omega}}_{\{1,2\}}\sigma_2 = \hat{\boldsymbol{\omega}}_{\{1,3\}}\sigma_1 \tag{53}$$

From equations (51), (52), (53), we can express σ_2 and σ_3 in terms of σ_1:

$$\sigma_2 = \frac{\hat{\boldsymbol{\omega}}_{\{1,3\}}}{\hat{\boldsymbol{\omega}}_{\{1,2\}}}\sigma_1 \quad \text{(from (53))} \tag{54}$$

$$\sigma_3 = \frac{\hat{\boldsymbol{\omega}}_{\{2,3\}}}{\hat{\boldsymbol{\omega}}_{\{1,2\}}}\sigma_1 \quad \text{(from (52))} \tag{55}$$

Substituting into equation (48):

$$\frac{\hat{\boldsymbol{\omega}}_{\{1,2\}}}{\sigma_1} + \frac{\hat{\boldsymbol{\omega}}_{\{1,3\}} \cdot \hat{\boldsymbol{\omega}}_{\{1,2\}}}{\hat{\boldsymbol{\omega}}_{\{1,3\}}\sigma_1} = n\bar{t}$$

$$\frac{\hat{\boldsymbol{\omega}}_{\{1,2\}} + \hat{\boldsymbol{\omega}}_{\{1,2\}}}{\sigma_1} = n\bar{t}$$

$$\sigma_1 = \frac{2\hat{\boldsymbol{\omega}}_{\{1,2\}}}{n\bar{t}} \tag{56}$$

Therefore:

$$\sigma_1 = \frac{2\hat{\boldsymbol{\omega}}_{\{1,2\}}}{n\bar{t}}, \quad \sigma_2 = \frac{2\hat{\boldsymbol{\omega}}_{\{1,3\}}}{n\bar{t}}, \quad \sigma_3 = \frac{2\hat{\boldsymbol{\omega}}_{\{2,3\}}}{n\bar{t}} \tag{57}$$

Now we recover the individual λ_j from the system:

$$\lambda_1 + \lambda_2 = \sigma_1 \tag{58}$$
$$\lambda_1 + \lambda_3 = \sigma_2 \tag{59}$$
$$\lambda_2 + \lambda_3 = \sigma_3 \tag{60}$$

Solving: $\lambda_1 = (\sigma_1 + \sigma_2 - \sigma_3)/2$, $\lambda_2 = (\sigma_1 - \sigma_2 + \sigma_3)/2$, $\lambda_3 = (-\sigma_1 + \sigma_2 + \sigma_3)/2$. Substituting the values of σ_i:

$$\hat{\lambda}_1 = \frac{1}{n\bar{t}}(\hat{\boldsymbol{\omega}}_{\{1,2\}} + \hat{\boldsymbol{\omega}}_{\{1,3\}} - \hat{\boldsymbol{\omega}}_{\{2,3\}}) \tag{61}$$

$$\hat{\lambda}_2 = \frac{1}{n\bar{t}}(\hat{\boldsymbol{\omega}}_{\{1,2\}} - \hat{\boldsymbol{\omega}}_{\{1,3\}} + \hat{\boldsymbol{\omega}}_{\{2,3\}}) \tag{62}$$

$$\hat{\lambda}_3 = \frac{1}{n\bar{t}}(-\hat{\boldsymbol{\omega}}_{\{1,2\}} + \hat{\boldsymbol{\omega}}_{\{1,3\}} + \hat{\boldsymbol{\omega}}_{\{2,3\}}) \tag{63}$$

This completes the derivation of equation (42). The general case follows the same pattern: pairwise subtraction of the m score equations reveals that all $(m-1)$-wise sums are proportional to their candidate set counts, yielding a linear system solvable in closed form. $\square$

Remark 5.3. The closed-form solution exists because the score equations reduce to a linear system in the $(m-1)$-wise sums. For $m = 3$, these are pairwise sums $\sigma_i = \lambda_j + \lambda_k$. The same algebraic structure persists for arbitrary m when $w = m-1$: pairwise subtraction of score equations reveals that all $(m-1)$-wise sums are proportional to their corresponding candidate set counts. However, for $w < m-1$, the score equations do not admit this linearization and closed-form solutions generally do not exist.

Remark 5.4 (Boundary Estimates). The closed-form MLE is the unconstrained solution to the score equations. For extreme candidate set distributions (e.g., $\hat{\boldsymbol{\omega}}_{\{2,3\}} > \hat{\boldsymbol{\omega}}_{\{1,2\}} + \hat{\boldsymbol{\omega}}_{\{1,3\}}$), the formula may yield $\hat{\lambda}_1 < 0$, which is outside the parameter space $(0, \infty)^m$.

Under the uniform masking model, $\mathrm{P}\{\hat{\lambda}_j < 0\} \to 0$ as $n \to \infty$ when $\lambda_j^\star > 0$. Negative estimates in finite samples suggest either:

1. Small sample size (insufficient averaging over candidate sets)

2. True parameter near zero (component rarely fails)

3. Potential model misspecification (uniform masking assumption violated)

When negative estimates occur, the constrained MLE (setting $\lambda_j = 0$ and re-solving for remaining components) also has closed form. For example, with $\lambda_1 = 0$:

$$\hat{\lambda}_2 = \frac{\hat{\boldsymbol{\omega}}_{\{1,2\}}}{(\hat{\boldsymbol{\omega}}_{\{1,2\}} + \hat{\boldsymbol{\omega}}_{\{1,3\}})\bar{t}}, \quad \hat{\lambda}_3 = \frac{\hat{\boldsymbol{\omega}}_{\{1,3\}}}{(\hat{\boldsymbol{\omega}}_{\{1,2\}} + \hat{\boldsymbol{\omega}}_{\{1,3\}})\bar{t}} \tag{64}$$

This corresponds to the two-component system where component 1 never fails.

The Fisher information matrix is:

$$\mathcal{I}(\boldsymbol{\lambda}^\star | w = 2) = \frac{1}{2(\lambda_1^\star + \lambda_2^\star + \lambda_3^\star)} \begin{bmatrix} \frac{1}{\lambda_1^\star + \lambda_2^\star} + \frac{1}{\lambda_1^\star + \lambda_3^\star} & \frac{1}{\lambda_1^\star + \lambda_2^\star} & \frac{1}{\lambda_1^\star + \lambda_3^\star} \\ \frac{1}{\lambda_1^\star + \lambda_2^\star} & \frac{1}{\lambda_1^\star + \lambda_2^\star} + \frac{1}{\lambda_2^\star + \lambda_3^\star} & \frac{1}{\lambda_2^\star + \lambda_3^\star} \\ \frac{1}{\lambda_1^\star + \lambda_3^\star} & \frac{1}{\lambda_2^\star + \lambda_3^\star} & \frac{1}{\lambda_1^\star + \lambda_3^\star} + \frac{1}{\lambda_2^\star + \lambda_3^\star} \end{bmatrix} \tag{65}$$

20

The inverse (asymptotic variance-covariance) is obtained by symbolic matrix inversion:

$$\mathcal{I}^{-1}(\boldsymbol{\lambda}^\star | w = 2) = (\lambda_1^\star + \lambda_2^\star + \lambda_3^\star) \begin{bmatrix} \lambda_1^\star + \lambda_2^\star + \lambda_3^\star & -\lambda_3^\star & -\lambda_2^\star \\ -\lambda_3^\star & \lambda_1^\star + \lambda_2^\star + \lambda_3^\star & -\lambda_1^\star \\ -\lambda_2^\star & -\lambda_1^\star & \lambda_1^\star + \lambda_2^\star + \lambda_3^\star \end{bmatrix} \tag{66}$$

This can be verified by direct multiplication: $\mathcal{I} \cdot \mathcal{I}^{-1} = \mathbf{I}_3$.

The asymptotic mean squared error (trace of variance-covariance) is:

$$\mathrm{MSE}(\hat{\boldsymbol{\lambda}}_n) = \frac{3(\lambda_1^\star + \lambda_2^\star + \lambda_3^\star)^2}{n} \tag{67}$$

5.2   Candidate Sets of Size One

When $w = 1$, each observation identifies the exact failed component. This represents the no-masking case. The MLE is:

$$\hat{\boldsymbol{\lambda}}_n = \frac{1}{n\bar{t}} \begin{pmatrix} \hat{\boldsymbol{\omega}}_{\{1\}} \\ \hat{\boldsymbol{\omega}}_{\{2\}} \\ \hat{\boldsymbol{\omega}}_{\{3\}} \end{pmatrix} \tag{68}$$

The Fisher information matrix is diagonal:

$$\mathcal{I}(\boldsymbol{\lambda}^\star | w = 1) = \frac{1}{\lambda_1^\star + \lambda_2^\star + \lambda_3^\star} \mathrm{diag}\left(\frac{1}{\lambda_1^\star}, \frac{1}{\lambda_2^\star}, \frac{1}{\lambda_3^\star} \right) \tag{69}$$

The asymptotic variance-covariance is:

$$\mathcal{I}^{-1}(\boldsymbol{\lambda}^\star | w = 1) = (\lambda_1^\star + \lambda_2^\star + \lambda_3^\star)\mathrm{diag}(\lambda_1^\star, \lambda_2^\star, \lambda_3^\star) \tag{70}$$

The MSE is:

$$\mathrm{MSE}(\hat{\boldsymbol{\lambda}}_n | w = 1) = \frac{(\lambda_1^\star + \lambda_2^\star + \lambda_3^\star)^2}{n} \tag{71}$$

which is exactly $1/3$ the MSE when $w = 2$, reflecting the additional information from exact component identification.

Remark 5.5. The $1/3$ MSE ratio holds regardless of the parameter values $\boldsymbol{\lambda}^\star$ because both MSE expressions factor as $(\sum_j \lambda_j^\star)^2/n$ times a constant that depends only on w, not on the individual $\lambda_j^\star$. This result is specific to the three-component system; for general m and w, the MSE ratio depends on both system dimension and masking cardinality.

5.3 Numerical Validation

We validate the asymptotic theory through Monte Carlo simulation studies. For each configuration, we generate r independent samples of size n, compute the MLE for each sample, and compare the empirical covariance with the theoretical asymptotic covariance.

5.3.1 Experimental Design

We consider multiple parameter configurations to assess robustness:

- **Symmetric**: $\boldsymbol{\lambda}^\star = (3, 3, 3)^\top$ (equal failure rates)

- **Moderate imbalance**: $\boldsymbol{\lambda}^\star = (2, 3, 4)^\top$ (1:1.5:2 ratio)

- **Strong imbalance**: $\boldsymbol{\lambda}^\star = (1, 3, 5)^\top$ (1:3:5 ratio)

For each configuration, we vary the sample size: $n \in \{100, 500, 1000, 5000\}$. All simulations use $r = 10{,}000$ Monte Carlo replications and masking cardinality $w = 2$.

5.3.2 Comparison Metrics

We quantify agreement between theoretical and empirical covariances using:

- **Frobenius norm of difference**: $\|\widehat{\mathrm{Cov}} - \frac{1}{n}\mathcal{I}^{-1}\|_F$

- **Maximum element-wise absolute difference**: $\max_{i,j} |\widehat{\mathrm{Cov}}_{ij} - \frac{1}{n}[\mathcal{I}^{-1}]_{ij}|$

- **Relative Frobenius error**: $\|\widehat{\mathrm{Cov}} - \frac{1}{n}\mathcal{I}^{-1}\|_F / \|\frac{1}{n}\mathcal{I}^{-1}\|_F$

5.3.3 Representative Results

Configuration 1: $\boldsymbol{\lambda}^\star = (2, 3, 4)^\top$, $n = 1000$, $w = 2$

The theoretical asymptotic variance-covariance matrix (evaluated at $n = 1000$) is:

$$\frac{1}{1000}\mathcal{I}^{-1}(\boldsymbol{\lambda}^\star | w = 2) = \begin{bmatrix} 0.081 & -0.036 & -0.027 \\ -0.036 & 0.081 & -0.018 \\ -0.027 & -0.018 & 0.081 \end{bmatrix} \tag{72}$$

The empirical variance-covariance from $r = 10{,}000$ Monte Carlo replications is:

$$\widehat{\text{Cov}} = \begin{bmatrix} 0.081 & -0.037 & -0.027 \\ -0.037 & 0.082 & -0.018 \\ -0.027 & -0.018 & 0.081 \end{bmatrix} \tag{73}$$

Comparison metrics:

- Frobenius norm error: 0.0012

- Maximum element-wise error: 0.001

- Relative Frobenius error: 0.79%

Configuration 2: $\boldsymbol{\lambda}^\star = (1, 3, 5)^\top$, $n = 5000$, $w = 2$

The theoretical and empirical covariances agree to within 0.3% relative Frobenius error, confirming that asymptotic approximations remain accurate even under strong parameter imbalance when n is sufficiently large.

5.3.4 Finite-Sample Behavior

To assess convergence rates, we plot the relative Frobenius error as a function of sample size for each parameter configuration. The error decreases approximately as $O(n^{-1/2})$, consistent with central limit theorem predictions.

For $n = 100$, the asymptotic approximation shows 3-5% relative error, which decreases to less than 1% for $n \geq 1000$. This suggests that asymptotic confidence intervals are reliable for practical sample sizes exceeding $n = 1000$.

5.3.5 Implementation Note

Detailed simulation code, including data generation, MLE computation via Corollary 5.2, and metric calculation, is available at https://github.com/queelius/series_system_estimation. The simulations were implemented in Python using NumPy for matrix operations and verified against independent R implementations.

6 Conclusion

We have developed a comprehensive framework for statistical inference in series systems with exponentially distributed component lifetimes when failure data is masked. The exponential case admits closed-form expressions for

the maximum likelihood estimator, Fisher information matrix, and sufficient statistics.

The main practical insights are:

1. The information content of masked data is quantified by the Fisher information matrix, which depends on the masking cardinality w

2. Minimal sufficient statistics are the mean system lifetime and candidate set frequencies

3. For m-component systems with $w = m - 1$, a closed-form MLE exists (the three-component case is developed in detail)

4. Asymptotic confidence intervals provide practical uncertainty quantification

6.1 Model Assumptions and Limitations

This work makes two key assumptions that limit its direct applicability but enable analytical tractability:

Uniform Masking Model: The assumption that non-failed components are equally likely to be included in the candidate set (Definition 2.3) is *not realistic in most practical scenarios*. Real diagnostic processes exhibit systematic biases:

- *Spatial clustering*: Components in the same physical location or subsystem are often grouped in candidate sets

- *Accessibility bias*: Easily inspected components may be overrepresented in candidate sets

- *Failure mode correlation*: Components with similar failure signatures are more likely to be co-nominated

- *Cost-based selection*: Expensive diagnostic tests may systematically exclude certain components

When the uniform masking assumption is violated, the likelihood in Section 3 is misspecified: the candidate-set term (7) is wrong. The resulting MLE converges to a *pseudo-true* parameter that solves the expected score under the incorrect masking model, which generally differs from the real

24

parameter unless the non-uniform masking probabilities happen to be proportional across components. Intuitively, non-uniform masking up- or down-weights certain components in the log-likelihood, so the estimator absorbs masking bias instead of recovering the true failure rates. The Fisher information matrix derived here thus gives efficiency and variance only under uniform masking; with non-uniform masking, both bias and variance depend on the (unmodeled) masking mechanism.

Exponential Lifetimes: The constant-hazard assumption is appropriate for:

- Random external shocks (e.g., power surges, environmental failures)

- Systems without aging or wear-out (e.g., early-life failures)

- Components with memoryless failure processes

The exponential assumption fails when:

- Components exhibit wear-out (increasing hazard rate)

- Infant mortality dominates (decreasing hazard rate)

- Failure mechanisms are time-dependent

For non-exponential distributions (Weibull, gamma, lognormal), the likelihood equations generally do not admit closed-form solutions, and the Fisher information matrix must be computed numerically or via simulation.

Justification: Despite these limitations, we adopt these assumptions because they enable *complete analytical solutions*. The resulting formulas provide:

1. Baseline performance bounds for assessing more complex models

2. Insight into the fundamental information structure of masked data

3. Computational efficiency for quick parameter estimation

4. A reference model analogous to the role of the normal distribution in statistics

Practitioners should validate the uniform masking assumption by examining diagnostic process documentation or performing goodness-of-fit tests on candidate set patterns. When the assumption is questionable, simulation studies can assess robustness or alternative models (e.g., component-specific masking probabilities) should be considered.

6.2 Extensions

Several extensions would broaden the applicability of this framework:

1. **Non-uniform masking models**: Relax the uniform masking assumption to allow component-specific inclusion probabilities α_j or conditional dependence structures. For example:

   - Accessibility-weighted masking: $\mathrm{P}\{j \in \mathrm{C} | \mathrm{K} = k, j \neq k\} = \alpha_j$ depends on component accessibility
   - Cluster-based masking: Components within the same subsystem have correlated inclusion probabilities
   - Parametric masking models with estimable diagnostic parameters

   Such models sacrifice analytical tractability but more accurately reflect real diagnostic processes. The uniform masking results provide baseline comparisons and limiting cases.

2. **Variable masking cardinality**: Remark 4.6 provides the optimal combination formula when w varies across observations. Further extensions could model w as random, depending on diagnostic difficulty or component type

3. **Non-exponential lifetime distributions**: Extend to Weibull (aging/wear-out), gamma (multi-stage failures), or lognormal (time-dependent hazards) distributions. Numerical methods will be required, but asymptotic theory remains applicable.

4. **Covariate effects**: Incorporate covariate information (operating conditions, environmental factors) via proportional hazards models or accelerated failure time frameworks

5. **Bayesian approaches**: When prior information about component reliabilities is available, derive posterior distributions for $\boldsymbol{\lambda}$ and posterior predictive distributions for future failures

6. **Optimal diagnostic design**: Given costs of diagnostic tests and benefits of precise failure identification, design inspection strategies that optimize the information-cost tradeoff

The uniform masking, exponential lifetime framework provides a foundation for understanding masked system data. While restrictive, the analytical tractability enables clear insight into the information structure that extends conceptually to more complex settings. The closed-form results serve as benchmarks for evaluating numerical methods and assessing the efficiency loss from model misspecification.

A   Numerical Solution Methods

For cases without closed-form solutions, the MLE must be computed numerically. The Newton-Raphson algorithm is effective:

Algorithm 1: Newton-Raphson for Exponential MLE

Input: Initial guess $\boldsymbol{\lambda}^{(0)}$, tolerance ϵ
Output: MLE $\hat{\boldsymbol{\lambda}}_n$
1 $k \leftarrow 0$;
2 **repeat**
3   Compute score $\boldsymbol{s}^{(k)} = \nabla\ell(\boldsymbol{\lambda}^{(k)}|\mathbf{M}_n)$;
4   Compute Hessian $\mathbf{H}^{(k)} = \nabla^2\ell(\boldsymbol{\lambda}^{(k)}|\mathbf{M}_n)$;
5   Update $\boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)} - [\mathbf{H}^{(k)}]^{-1}\boldsymbol{s}^{(k)}$;
6      $k \leftarrow k + 1$;
7 **until** $\|\boldsymbol{\lambda}^{(k+1)} - \boldsymbol{\lambda}^{(k)}\| < \epsilon$;
8 **return** $\boldsymbol{\lambda}^{(k)}$

The Hessian for the exponential log-likelihood is:

$$\left[\nabla^2\ell(\boldsymbol{\lambda}|\mathbf{M}_n)\right]_{jk} = -\sum_{i=1}^{n} \frac{\mathbb{1}_{\mathrm{C}_i \times \mathrm{C}_i}(j,k)}{\left(\sum_{p \in \mathrm{C}_i} \lambda_p\right)^2} \tag{74}$$

Convergence is typically rapid when initialized at a reasonable starting point such as $\boldsymbol{\lambda}^{(0)} = (1/\bar{t}, \ldots, 1/\bar{t})$.

References

[1] Marcus A. Agustin. Systems in series. In *Wiley Encyclopedia of Operations Research and Management Science*. John Wiley & Sons, 2010.

[2] R. E. Barlow and F. Proschan. *Statistical Theory of Reliability and Life Testing: Probability Models*. Holt, Rinehart and Winston, New York, 1975.

[3] Peter Bickel and Kjell Doksum. *Mathematical Statistics*, volume 1, chapter 2, page 117. Prentice Hall, 2 edition, 2000.

[4] D. R. Cox. The analysis of exponentially distributed life-times with two types of failure. *Journal of the Royal Statistical Society: Series B (Methodological)*, 21(2):411–421, 1959.

[5] G. E. Dinse. Nonparametric estimation for partially-complete time and type of failure data. *Biometrics*, 38(2):417–431, 1982.

[6] B. J. Flehinger, B. Reiser, and E. Yashchin. Survival with competing risks and masked causes of failures. *Biometrika*, 85(1):151–164, 1998.

[7] B. J. Flehinger, B. Reiser, and E. Yashchin. Parametric modeling for survival with competing risks and masked failure causes. *Lifetime Data Analysis*, 8(2):177–203, 2002.

[8] Huairui Guo, Pengying Niu, and Ferenc Szidarovszky. Estimating component reliabilities from incomplete system failure data. In *Proceedings of the Annual Reliability and Maintainability Symposium (RAMS)*, pages 1–6. IEEE, 2013.

[9] I. Guttman, D. K. J. Lin, B. Reiser, and J. S. Usher. Dependent masking and system life data analysis: Bayesian inference for two-component systems. *Lifetime Data Analysis*, 1(1):87–100, 1995.

[10] Lynn Kuo and Tao Y. Yang. Masked failure data: Bayesian modeling. In *Encyclopedia of Statistics in Quality and Reliability*. John Wiley & Sons, 2007.

[11] D. K. J. Lin, J. S. Usher, and F. M. Guess. Exact maximum likelihood estimation using masked system data. *IEEE Transactions on Reliability*, 42(4):631–635, 1993.

[12] M. Miyakawa. Analysis of incomplete data in competing risks model. *IEEE Transactions on Reliability*, 33(4):293–296, 1984.

[13] B. Reiser, I. Guttman, D. K. J. Lin, J. S. Usher, and F. M. Guess. Bayesian inference for masked system lifetime data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 44(1):79–90, 1995.

[14] A. M. Sarhan. Reliability estimations of components from masked system life data. *Reliability Engineering & System Safety*, 74(1):107–113, 2001.

[15] John S. Usher and T. J. Hodgson. Maximum likelihood analysis of component reliability using masked system life-test data. *IEEE Transactions on Reliability*, 37(5):550–555, 1988.