Integrated Likelihood Inference in Poisson Distributions

Timothy Ruel
Department of Statistics and Data Science, Northwestern University

August 31, 2024

Abstract

The text of your abstract. 200 or fewer words.

Keywords: Directly standardized rate, Integrated likelihood ratio statistic, Maximum integrated likelihood estimator, Profile likelihood, Weighted sum, Zero score expectation parameter

1 Introduction

Consider a vector $\theta = (\theta_1, ..., \theta_n)$ in which each component represents the mean of a distinct Poisson process. We may obtain a sample of values from each process through repeated measurements of the number of events it generates over a fixed period of time. Suppose we have done so, and let X_{ij} represent the jth count from the ith sample, so that $X_{ij} \sim \text{Poisson}(\theta_i)$ for $i = 1, ..., n$ and $j = 1, ..., m_i$. The purpose of this paper is to consider likelihood-based inference for a real-valued parameter of interest $\psi = g(\theta)$, where $\theta \in \Theta \subset \mathbb{R}_+^m$ is unknown and $g : \Theta \to \Psi$ is a known twice continuously differentiable function.

The log-likelihood function of the model is given by

$$\ell(\theta) = \sum_{i=1}^{n} \left[X_i \log(\theta_i) - \theta_i\right]. \tag{1}$$

Note that while ℓ is a function of the full m-dimensional vector parameter θ, ψ is just a scalar. This decrease in dimension induces a nuisance parameter λ in the model that typically must be eliminated from the log-likelihood function before inference regarding ψ can be conducted.[1] Furthermore, ψ need not explicitly be equal to one of the components of θ but instead may be defined as any function g of θ satisfying the requirements mentioned above. We refer to ψ and λ as being *implicit* parameters in such cases. Note that in general a closed form expression for an implicit nuisance parameter will not exist.

The standard procedure for eliminating λ from the log-likelihood function involves choosing some method with which to summarize $\ell(\theta)$ over its possible values while holding ψ fixed in place. This effectively reduces $\ell(\theta)$ to a simpler function depending on ψ alone,

[1] Intuitively, we can think of λ as representing the portion of θ that remains in the other $m-1$ dimensions of Θ that are not occupied by ψ.

having replaced each dimension of θ that depends on λ with a static summary of the values in its parameter space. We call this new function a pseudo-log-likelihood function for ψ and denote its generic form as $\ell(\psi)$. As we encounter specific types of pseudo-log-likelihoods, we will introduce more specialized notation as needed. Note that while it usually has properties resembling one, $\ell(\psi)$ is not itself considered a genuine log-likelihood function, and there will always be some degree of information contained within the data lost as a result of the nuisance parameter's elimination.

Perhaps the most straightforward method of summarization we can use to construct $\ell(\psi)$ is to maximize $\ell(\theta)$ over all possible of values of θ for a fixed value of ψ. This yields what is known as the *profile* log-likelihood function, formally defined as

$$\ell_p(\psi) = \sup_{\theta \in \Theta:\, g(\theta) = \psi} \ell(\theta). \tag{2}$$

In the case where an explicit nuisance parameter exists, Equation 2 is equivalent to replacing λ with its conditional maximum likelihood estimate given ψ:

$$\ell_p(\psi) = \ell(\psi, \hat{\lambda}_\psi). \tag{3}$$

2 Integrated Likelihood Functions

3 Application to Poisson Models

We now turn our attention to the task of using the ZSE parameterization to construct an integrated likelihood that can be used to make inferences regarding a parameter of interest derived from the Poisson model described in the introduction. We will

4 Inference for the Weighted Sum of Poisson Means

Consider a set of Poisson means corresponding to n distinct populations, where n is a known positive integer. Let θ_i denote the mean of population i, and suppose we have drawn an independent sample of size m_i from each population, where X_{ij} represents the jth count from the ith sample. Then we have

$$X_{ij} \sim \text{Poisson}(\theta_i), \;\; i = 1, ..., n; \;\; j = 1, ..., m_i.$$

Consider the weighted sum

$$Y = \sum_{i=1}^{n} w_i X_i,$$

where each w_i is a known constant greater than zero. Suppose we take for our parameter of interest the expected value of this weighted sum, so that

$$\psi \equiv \mathrm{E}(Y) = \sum_{i=1}^{n} w_i \theta_i.$$

The maximum likelihood estimate (MLE) for θ_i is simply the sample mean $\hat{\theta}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} X_{ij}$.

Examples