We are motivated by problems that arise in a number of applications such as Online Marketing and Explosives detection, where the observations are usually modeled using Poisson statistics. We model each observation as a Poisson random variable whose mean is a sparse linear superposition of known patterns. Unlike many conventional problems observations here are not identically distributed since they are associated with different sensing modalities. We analyze the performance of a maximum likelihood (ML) decoder, which for our Poisson setting involves a non-linear optimization but yet is computationally tractable. We derive fundamental sample complexity bounds for sparse recovery when the measurements are contaminated with Poisson noise. In contrast to the least-squares linear regression setting with Gaussian noise, we observe that in addition to sparsity, the scale of the parameters also fundamentally impacts l2 error in the Poisson setting. We show that our upper bounds are tight under suitable regularity conditions. Specifically, we derive a minimax matching lower bound on the mean-squared error and show that our constrained ML decoder is minimax optimal for this regime.