Addressing over/under dispersion in a Poisson observation model

M. Rule

August 30, 2017

The neural-field Cox-process model currently assumes a Poisson observation model. However, firing may be periodic (under-dispersed) or bursty (over-dispersed) and deviate from the Poisson model.

One way to handle this is to add a multiplicative parameter to the spiking observations to capture over/under dispersion. The resulting log-likelihood resembles that for Poisson observations, but doesn't correspond to a proper distribution over discrete spike-counts. It is sometimes called the *quasi-likelihood* approach.

0.1 Quasi-likelihood for Poisson count data

In the non-spatial system, the Poisson observation model for counts y given intensity λ is

$$\Pr(y|\lambda) = \frac{1}{y!}\lambda^y e^{-\lambda}$$

One way to handle over/under dispersion is to pretend that the Poisson distribution is a continuous distribution over the positive real numbers (as opposed to positive integers), and to add a scale parameter to this distribution.

The quasi-likelihood approach introduces a parameter α, that scales both the spike counts and the rates.

$$\Pr(y|\lambda) \propto \frac{(\alpha\lambda)^{\alpha y} e^{-\alpha\lambda}}{\Gamma(\alpha y + 1)}$$

The log-probability of y given λ is then

$$\ln \Pr(y|\lambda) = (\alpha y)\ln(\alpha\lambda) - \alpha\lambda - \ln\Gamma(\alpha y + 1) + \text{constant}$$

Typically, we're given (fixed) observations y, and are only interested in optimizing the log-likelihood up to a constant, so one might write:

$$\ln \Pr(y|\lambda) = (\alpha y)\ln(\alpha\lambda) - \alpha\lambda + \text{constant}$$

The quasi-likelihood can be optimized using the same code as one would use for the GLM. In this case, α can be folded in to a gain adjustment on the rates λ, and the count data is pre-multiplied by α before passing them on to the GLM inference routine.

0.2 In a spatially extended system with a linear observation model

In the spatially extended case, $\lambda(x)$ depends on spatial coordinates $ x$. Spikes occur as a spatial Poisson process

$$y(x) = \sum_i \delta(x - x_i)$$

consisting of a sum of delta distributions at various locations. For the Poisson process, the log probability is

$$\ln \Pr(y(x)) = \int_x y(x) \ln \lambda(x) - \int_x \lambda(x) + \text{constant}$$

In our model, we have a latent variable $A(x)$ that drives retinal ganglion cell spiking at rate $\lambda = \gamma_0 A(x) + \beta_0$.

We can extend this model to account for over and under-dispersion is to add a scalar multiplier α to the observed counts y, so that the measurements enter as $\alpha y(x)$. We implicitly fold α into the gain $\gamma \leftarrow \alpha\gamma_0$ and bias $\beta \leftarrow \alpha\beta_0$ parameters.

The spatially-extended log-probability in terms of latent field $A(x)$, with added parameters reflecting a gain, bias, and dispersion, is

$$\ln \Pr(y|\lambda) = \int_x \alpha y(x) \ln[\gamma A(x) + \beta] - \int_x [\gamma A(x) + \beta] + \text{constant}$$

For inference, we are interested in optimizing latent states $A(x)$ given a fixed observation $y(x)$, so terms not affecting $A(x)$ can be dropped:

$$\mathcal{L} = \ln \Pr(y|\lambda) = \alpha \int_x y(x) \ln[A(x) + \beta/\gamma] - \gamma \int_x A(x) + \text{constant}$$

The measurement update is performed using a Gaussian prior, and computed using the Laplace approximation for which one needs the first and second derivatives of the likelihood (quasi-likelihood) in $A(x)$.

$$\frac{\partial \mathcal{L}}{\partial x_1} = -\gamma + \frac{\alpha y(x_1)}{A(x_1) + \beta/\gamma}$$

$$\frac{\partial^2 \mathcal{L}}{\partial x_1 \partial x_2} = -\delta(x_1 - x_2) \frac{\alpha y(x_1)}{[A(x_1) + \beta/\gamma]^2}$$

These gradients are identical to those of the linear $\lambda(x) = \gamma A(x) + \beta$ observation model. The parameter α can be incorporated without modifying the measurement update code by pre-multiplying the counts $y(x)$ by α.