Local Poisson regression

1. Bandwidth choice for the local Poisson regression

Modify the functions h.cv.sm.binomial and loglik.CV to obtain a bandwidth choice method for the local Poisson regression based on the leave-one-out cross-validation (loo-CV) estimation of the expected likelihood of an independent observation.

Remember that the loo-CV estimation of the expected log-likelihood of an independent observation, when using h as bandwidth, is

$$\ell_{CV}(h) = \frac{1}{n} \sum_{i=1}^{n} \log \left(\widehat{\Pr}_{h}^{(-i)} (Y = y_i | X = x_i) \right),$$

where $\widehat{\Pr}_h^{(-i)}(Y=y_i|X=x_i)$ is an estimation of

$$\Pr(Y = y_i | X = x_i) = e^{-\lambda_i} \frac{\lambda_i^{y_i}}{y_i!},$$

and

$$\lambda_i = \mathbb{E}(Y|X=x_i)$$

should be estimated by maximum local likelihood using h as bandwidth (for instance, using the function $\mathtt{sm.poisson}$ from the R package \mathtt{sm}).

2. Local Poisson regression for Country Development Data

Consider the country development dataset (file HDI.2017.subset.csv) containing information on development indicators measured in 179 countries (Source: [Human Development Data (1990-2017)](http://hdr.undp.org/en/data), The Human Development Report Office, United Nations). Variable le.fm always takes non-negative values. Define le.fm.r as the rounded value of le.fm: le.fm.r <- round(le.fm)

Fit a local Poisson regression modeling le.fm.r as a function of Life.expec. Use sm.poisson from the R package sm with the bandwidth obtained by loo-CV.