

③

Replicated data

$$\text{prob}(y) = \sum_{n=N}^{\infty} P_0(n; \lambda) \prod_{i=1}^d \text{Bin}(y_i; n, p)$$

$$N = \max(y)$$

Using recursive formulas:

$$\begin{aligned} & \cancel{P_0(n; \lambda)} \prod_{i=1}^d \text{Bin}(y_i; n, p) \\ &= P_0(n+1; \lambda) \prod_{i=1}^d \text{Bin}(y_i; n+1, p) \\ &= \frac{(1-p)^d \lambda}{n} \prod_{i=1}^d \frac{n}{n-y_i} \end{aligned}$$

so its

$$\text{prob}(y) = P_0(N; \lambda) \prod_{i=1}^d \text{Bin}(y_i; N, p)$$

• fac

where fac is computed as

$$\text{fac} = 1$$

for nn in nmax: (y_{max}+1)

$$\text{fac} = \cancel{1} + \text{fac} \cdot \frac{\lambda (1-p)^d}{nn} \prod_{i=1}^d \frac{nn}{nn - y_i}$$

①

Jan 17th-18

N-mix model : details

Model

$$y \sim \text{Bin}(n, p)$$

$$n \sim \text{Pois}(\lambda)$$

Normally, $p = p(x)$, while, $\log(\lambda) = x^T \beta$

The likelihood is

$$\text{prob}(y) = \sum_{n=y}^{\infty} \text{Pois}(n; \lambda) \cdot \text{Bin}(y; n, p)$$

There is a nice recursive formula for this density, using these

$$\text{Pois}(n; \lambda) = \text{Pois}(n-1; \lambda) \frac{\lambda}{n}$$

$$\text{Bin}(y; n, p) = \text{Bin}(y; n-1, p) \frac{n}{n-y} (1-p)$$

So that

$$\begin{aligned} \text{Pois}(n; \lambda) \cdot \text{Bin}(y; n, p) &= \text{Pois}(n-1; \lambda) \text{Bin}(y; n-1, p) \\ &\quad \cdot \frac{\lambda}{n-y} (1-p) \end{aligned}$$

Let $f_i \equiv \frac{\lambda(1-p)}{i}$, $i = 1, 2, \dots, n$ then

then

2

$$\begin{aligned}
 L &= \sum_{n=y}^{n_{\max}} \text{Pois}(n; \lambda) \cdot \text{Bin}(y; n, p) \\
 &= \text{Pois}(y; \lambda) \cdot \text{Bin}(y; y, p) \cdot \left\{ \begin{aligned} &1 + f_1 + f_1 f_2 + f_1 f_2 f_3 + \dots + f_1 \dots f_{n_{\max}} \end{aligned} \right\} \\
 &= \text{Pois} \cdot \text{Bin} \left\{ \underbrace{1 + f_1 (1 + f_2 (1 + f_3 (1 + f_4 \dots)))}_{f_{\text{acc}}} \right\}
 \end{aligned}$$

~~acc~~ can be computed,

$$f_{\text{acc}} = 1$$

for i in $1:n_{\max}$:

$$f_{\text{acc}} = 1 + f_{\text{acc}} \cdot \frac{\lambda(1-p)}{i}$$

then

$$L = \text{Pois}(y; \lambda) \cdot \text{Bin}(y; y, p) \cdot f_{\text{acc}}$$

$$n_{\max} \text{ could be } \left\lceil \frac{\lambda(1-p)}{\epsilon} \right\rceil$$

$$\text{so that } \frac{\lambda(1-p)}{n_{\max}} \leq \epsilon$$

(4)

Using negative binomial instead of Poisson.
In this case then

$$n \sim \text{nbinom}(\lambda, \delta)$$

where λ is the mean ~~exp~~ and δ is the dispersion parameter, following the notation in dnbinom in R. The density is

$$f(n) = \frac{\Gamma(n+\delta)}{\Gamma(\delta) n!} q^\delta (1-q)^n$$

$$\text{where } q = p/(1+p) = \frac{\delta}{\delta + \lambda}$$

Also this has a recursive formulation, as

$$\begin{aligned} f(n) &= \frac{(n+\delta-1) \Gamma(n+\delta-1)}{\Gamma(\delta) n(n-1)!} q^\delta (1-q)^{n-1} (1-q) \\ &= f(n-1) \cdot \frac{n+\delta-1}{n} (1-q) \end{aligned}$$

This means that

$$\text{nbinom}(n; \lambda, \delta) \cdot \text{Bin}(y; n, p)$$

$$= \text{nbinom}(n-1; \lambda, \delta) \text{Bin}(y; n-1, p)$$

$$\begin{aligned} &\cdot \frac{n+\delta-1}{n} (1-q) \cdot \frac{n}{n-y} (1-p) \\ &\rightarrow = \frac{n+\delta-1}{n-y} (1-q)(1-p) \end{aligned}$$

Again, following the same logic, we can compute the likelihood as.

$$n \text{ binom}(y; \lambda, \delta) \cdot \text{Bin}(y; q, p) \cdot f$$

where f is computed as.

$$f = 1$$

for n in $n_{\max}:(y+1)$

$$f = 1 + f \cdot \frac{n + \delta - 1}{n - y} (1 - q)(1 - p)$$

For replicated data, we get.

$$\prod_{i=1}^d \{ n \text{ binom}(N; \lambda, \delta) \cdot \prod_{i=1}^d \text{Bin}(y_i; q, p) \} \cdot f.$$

$$f = 1$$

for n in $n_{\max}:(y_{\max} + 1)$

$$f = 1 + f \cdot \frac{n + \delta - 1}{n} (1 - q)(1 - p)^d \prod_{i=1}^d \frac{n}{n - y_i}$$