(3) Replicated data

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

$$N = \max(y)$$

Using recursive formulas:

$$P_o(n;\lambda) \prod_{i=1}^{d} Bin(y_i; n, p)$$

$$= P_o(n-1;\lambda) \prod_i Bin(y_i; n-1, p)$$

$$\cdot \frac{(1-p)^d \lambda}{n} \prod_{i=1}^{d} \frac{n}{n-y_i}$$

so its

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

$$\cdot fac$$

where facc is computed as

fac = 1

~~for~~ nn in nmax : (ymax+1)

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

1

Jan 17th-17

N-mix model : details

Model $y \sim Bin(n, p)$
$n \sim Pois(\lambda)$

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

The likelihood is

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

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

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

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

So that

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

Let $f_i = \frac{\lambda(1-p)}{i}$, $i = 1, 2, \ldots n_{max}$

then

(2)

$$L = \sum_{n=y}^{n_{max}} Pois(n; \lambda) \cdot Bin(y; n, p)$$

$$= Pois(y; \lambda) \cdot Bin(y; y, p) \cdot \{ 1 + f_1 + f_1 f_2 + f_1 f_2 f_3 + \dots + f_1 \cdots f_{n_{max}} \}$$

$$= Pois \cdot Bin \underbrace{\{ 1 + f_1 (1 + f_2 (1 + f_3 (1 + f_4 (1 + \dots))))\}}_{fac.}$$

~~fac~~ fac can be computed..

```
fac = 1
for i in imax:1
        fac = 1 + fac * λ(1-p)/i
```

then

$$\underline{L = Pois(y; \lambda) \cdot Bin(y; y, p) \cdot fac.}$$

imax could be $\left\lfloor \frac{\lambda(1-p)}{\varepsilon} \right\rfloor$

so that $\frac{\lambda(1-p)}{imax} \approx \varepsilon$

Using negative binomial instead ~~of~~ Poisson.
In this case then

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

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

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

where $q = \text{prob} = \frac{\delta}{\delta+\lambda}$

Also this has a recursive formulation, as

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

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

This means that

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

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

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

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

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

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

where f is computed as.

$f = 1$
for n in $nmax : (y+1)$

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

For replicated data, we get.

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

$f = 1$
for n in $nmax : (ymax+1)$

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