Negative Binomial PMF:

Y is counting y failures given r successes. Total # of trials $n = y + r$

$$f(y; p, r) = \binom{y+r-1}{r-1} p^r (1-p)^y$$

where p is the probability of success in each trial

Lemma: $$\binom{y+r-1}{r-1} = \binom{y+r-1}{y}$$

proof:

$$\binom{y+r-1}{r-1} = \frac{(y+r-1)!}{(r-1)!\,((y+r-1)-r-1)!} = \frac{(y+r-1)!}{(r-1)!\,y!}$$

$$= \frac{(y+r-1)!}{y!\,(r-1)!} = \frac{(y+r-1)!}{y!\,(y+r-1-y)!}$$

$$= \binom{y+r-1}{y}$$

$\square$

$2(\tilde{l} - l)$

$\overline{2l} - \underbrace{2l}_{\text{Dev. resid}}$

$- s\text{-}ll$

Therefore

$$f(y; p, r) = \binom{y+r-1}{y} p^r (1-p)^y$$

Mean: $\frac{r(1-p)}{p} =: \mu$, Variance: $\frac{r(1-p)}{p^2} = Var(Y)$

$\Rightarrow \mu p = r - rp$

$\Rightarrow p(r+\mu) = r$

$\Rightarrow p = r/(r+\mu) \Rightarrow Var(Y) = \mu / r/(r+\mu) = \mu/r\,(r+\mu) = \mu + \frac{\mu^2}{r}$

$$f(y; \mu, r) = \binom{y+r-1}{y} \left(\frac{r}{r+\mu}\right)^r \left(\frac{\mu}{r+\mu}\right)^y$$

$$= \frac{\Gamma(y+r)}{y!\,\Gamma(r)} \left(\frac{r}{r+\mu}\right)^r \left(\frac{\mu}{r+\mu}\right)^y$$

$\Gamma(n) = (n-1)!$
if $n \in \mathbb{N}, n \geq 1$
if n is a real # the $\Gamma(\cdot)$

is well defined

Define $r := 1/\alpha$, $\Rightarrow p = \frac{1}{\alpha}\left(\frac{1}{1/\alpha} + \mu\right) = 1/(1+\alpha\mu)$

$$f(y;\mu,\alpha) = \frac{\Gamma(y+1/\alpha)}{\Gamma(y+1)\,\Gamma(1/\alpha)} \left(\frac{1}{1+\alpha\mu}\right)^{1/\alpha} \left(\frac{\alpha\mu}{1+\alpha\mu}\right)^{y}$$

$\Rightarrow$ The single data log-likelihood is

$$l_i^{NB} = y_i \log\left(\frac{\alpha\mu_i}{1+\alpha\mu_i}\right) - \frac{1}{\alpha}\log(1+\alpha\mu_i) + \log\Gamma(y_i + 1/\alpha)$$
$$- \log\Gamma(y_i+1) - \log\Gamma(1/\alpha) \quad \cdots \quad (1)$$

Zero-truncated NB.

The probability of a zero count is

$$f(0;\mu,\alpha) = \frac{\Gamma(1/\alpha)}{\Gamma(1)\,\Gamma(1/\alpha)}\left(\frac{1}{1+\alpha\mu}\right)^{1/\alpha} = \left(\frac{1}{1+\alpha\mu}\right)^{1/\alpha}$$

The zero-truncated density is

$$f(y \mid y>0, \mu,\alpha) = \frac{f(y;\mu,\alpha)}{1-f(0,\mu,\alpha)}$$

$$= \frac{\dfrac{\Gamma(y+1/\alpha)}{\Gamma(y+1)\,\Gamma(1/\alpha)}\left(\dfrac{1}{1+\alpha\mu}\right)^{1/\alpha}\left(\dfrac{\alpha\mu}{1+\alpha\mu}\right)^{y}}{1-\left(1/(1+\alpha\mu)\right)^{1/\alpha}}$$

This gives us the single data log-likelihood;

$$l_i^{ZT} = l_i^{NB} - \log\left\{1-(1+\alpha\mu_i)^{-1/\alpha}\right\} \quad \cdots \quad (2)$$

Where (2) is the NB single data log-likelihood as given in (1)

Zero Inflated Negative Binomial (Hurdle model in Tutz (2012))

We use a similar parameterization as in Hood et al. (2017) SA I. ${}_{\text{erated}}^{\text{by a}}$

$$g(y;\mu,\alpha) = \begin{cases} 1-q & y=0 \\ q\, f(y \mid y>0, \mu, \alpha) & \text{otherwise} \end{cases}$$

where q is the probability of a positive counts

Similar to Hood et al. (2017), we adopt the unconstrained parameterization

$$\gamma = \log \mu \quad , \quad \eta = \log\{-\log(1-q)\} .$$

This gives us the single data log-likelihood (we drop the data index i)

$$l^{ZI} = \begin{cases} \log(1-q) & y=0 \\ \log q + \log f(y \mid y>0, \mu, \alpha) & \text{otherwise} \end{cases}$$

$$= \begin{cases} -e^{\eta} & y=0 \\ \log\left(1-e^{-e^{\eta}}\right) + l^{ZT} & \text{otherwise} \end{cases}$$

$$= \begin{cases} -e^{\eta} & y=0 \\ \log\left(1-e^{-e^{\eta}}\right) + l^{NB} - \log\left\{1-\left(1+\alpha e^{\gamma}\right)^{-1/\alpha}\right\} & \text{otherwise} \end{cases}$$

$$= \begin{cases} -e^{\eta} & y=0 \\ \log\left(1-e^{-e^{\eta}}\right) + l^{ZT} & \text{otherwise} \end{cases}$$

where $l^{ZT} = l^{NB} - \log\{1-(1+\alpha e^{\gamma})^{-1/\alpha}\}$ is the ZTNB log-likelihood.

and $l^{NB} = y \log\left(\alpha e^{\gamma}/_{1+\alpha e^{\gamma}}\right) - \frac{1}{\alpha}\log(1+\alpha e^{\gamma}) + \log\Gamma(y+1/\alpha)$

$$- \log\Gamma(y+1) - \log\Gamma(1/\alpha)$$

$$\Rightarrow l^{ZT} = y\log\left(\alpha e^{\gamma}/_{1+\alpha e^{\gamma}}\right) + \log\Gamma(y+1/\alpha) - \log\Gamma(y+1) - \log\Gamma(1/\alpha)$$
$$+ \log\left\{\frac{1}{(1+\alpha e^{\gamma})^{1/\alpha}}\right\} + \log\left\{\frac{1}{1-\frac{1}{(1+\alpha e^{\gamma})^{1/\alpha}}}\right\}$$

looking at the last two terms basically is

$$\log\left\{\frac{1}{(1+\alpha e^{\gamma})^{1/\alpha}-1}\right\} = -\log\{(1+\alpha e^{\gamma})^{1/\alpha}-1\}$$

$$\Rightarrow l^{ZT} = y\log\left\{\alpha e^{\gamma}/_{1+\alpha e^{\gamma}}\right\} + \log\Gamma(y+1/\alpha) - \log\Gamma(y+1)$$
$$- \log\Gamma(1/\alpha) - \log\{(1+\alpha e^{\gamma})^{1/\alpha}-1\}$$

$$= y\log\alpha + y\gamma - y\log(1+\alpha e^{\gamma}) - \log\{(1+\alpha e^{\gamma})^{1/\alpha}-1\}$$
$$+ \log\Gamma(y+1/\alpha) - \log\Gamma(y+1) - \log\Gamma(1/\alpha)$$

Here as well, some care is required to evaluate l^{ZT} without unnecessary overflow, since it is easy for $1-e^{-e^{\eta}}$ and $(1+\alpha e^{\gamma})^{1/\alpha}-1$ to evaluate to zero in finite precision arithmetic. Hence the limiting results $\log(1-e^{-e^{\eta}}) \to \log(e^{\eta}-e^{2\eta}+e^{3\eta}/6) \to \eta$ as $\eta\to-\infty$. Similar as $\gamma\to-\infty$, $\log\{(1+\alpha e^{\gamma})^{1/\alpha}-1\} \to \gamma$ and as $\alpha\to 0$ $\log\{(1+\alpha e^{\gamma})^{1/\alpha}-1\} \to \log(e^{e^{\gamma}}-1) \to e^{\gamma}$ provided $\gamma \not\to -\infty$

Implemented taking α as fixed.

Proof:

① $\log(1-e^{-e^{\eta}}) \to \log(e^{\eta}-e^{2\eta}/2+e^{3\eta}/6) \to \eta$ $n\to-\infty$

as $\eta\to-\infty$, $-e^{\eta}\to 0$,

we know that $e^{x} \to 1+x+\frac{x^2}{2}+\frac{x^3}{6}$ as $x\to 0$

$$\Rightarrow e^{-e^{\eta}} \to 1-e^{\eta}+e^{2\eta}/2-e^{3\eta}/6$$

$e^x := \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n$

$\Rightarrow$ as $\eta \to -\infty$, $\log(1-e^{-e^{\eta}})$

$$\to \log\left(1 - 1 + e^{\eta} - \frac{e^{2\eta}}{2} + \frac{e^{3\eta}}{6}\right)$$

$$= \log\left(e^{\eta} - \frac{e^{2\eta}}{2} + \frac{e^{3\eta}}{6}\right)$$

as $\eta \to -\infty$ $e^{2\eta}$ and $e^{3\eta} \to 0$ faster than e^{η}

$$\Rightarrow \log\left(e^{\eta} - e^{2\eta}/2 + \frac{e^{3\eta}}{6}\right)$$

$$\to \log e^{\eta} \to \eta$$

② $\log\left\{(1+\alpha e^{\gamma})^{1/\alpha} - 1\right\} \to \gamma$ as $\gamma \to -\infty$

as $\gamma \to -\infty$, $e^{\gamma} \to 0$ $\Rightarrow \alpha e^{\gamma} \to 0$ (fixed α)

We know that $(1+x)^{1/\alpha} \to 1 + \frac{1}{\alpha}x$ as $x \to 0$

$$\Rightarrow (1+\alpha e^{\gamma})^{1/\alpha} \to 1 + \frac{1}{\alpha}(\alpha e^{\gamma}) = 1 + e^{\gamma}$$

$$\Rightarrow \log\left\{(1+\alpha e^{\gamma})^{1/\alpha} - 1\right\} \to \log\left\{(1+e^{\gamma}) - 1\right\}$$

$$= \log(e^{\gamma}) = \gamma$$

as $\gamma \to -\infty$

③ $\log\left\{(1+\alpha e^{\gamma})^{1/\alpha} - 1\right\} \to e^{\gamma}$ as $\alpha \to 0$ for fixed γ

remember that $(1+\alpha e^{\gamma})^{1/\alpha} = \left(1 + \frac{e^{\gamma}}{r}\right)^{r}$,

So as $\alpha \to 0$, $r \to \infty$

and $\lim_{r\to\infty}\left(1+\frac{e^{\gamma}}{r}\right)^{r} = e^{e^{\gamma}}$ for $r \in \mathbb{N}_0$

$$\Rightarrow (1+\alpha e^{\gamma})^{1/\alpha} \to e^{e^{\gamma}} \text{ as } \alpha \to 0$$

$$\Rightarrow \log\left\{(1+\alpha e^{\gamma})^{1/\alpha} - 1\right\} \to \log\left\{e^{e^{\gamma}} - 1\right\} \text{ as } \alpha \to 0$$

also for very large γ, $e^{e^{\gamma}} - 1$ is dominated by $e^{e^{\gamma}}$

$$\Rightarrow \log\left\{e^{e^{\gamma}} - 1\right\} \to \log e^{e^{\gamma}} = e^{\gamma} \quad \square$$

On the expectation of y in the ZINB.

Recall the model's density

$$g(y;\mu,\alpha) = \begin{cases} 1-q & y=0 \\ q\, f(y \mid y>0, \mu, \kappa) & y>0 \end{cases}$$

where $f(y \mid y>0, \mu, \kappa) = \dfrac{f(y;\mu,\kappa)}{1-f(0;\mu,\alpha)}$

Therefore, the model adopts the form of hurdle models discussed in Tuty, section 7.8.

Let $v := \dfrac{P(y>0)}{1-f(0;\mu,\kappa)} \overset{\text{ZINB}}{=} q\,\tau$, since in ZINB $P(y>0) = q$ and $f(0;\mu,\kappa) = \dfrac{1}{(1+\alpha\mu)^{1/\alpha}}$, $\mu = e^{\gamma}$

Therefore, the mean is given by

$$\mathbb{E}[y] = \sum_{s=1}^{\infty} s\, f(s;\mu,\kappa)\, v = \sum_{i=1}^{\infty} s\, f(s;\mu,\kappa) \cdot \underbrace{\frac{\mathbb{P}(y>0)}{1-f(0;\mu,\kappa)}}_{\text{doesn't depend on the value of } s}$$

$$= \mathbb{P}(y>0) \sum_{s=1}^{\infty} s \frac{f(s;\mu,\kappa)}{1-f(0;\mu,\kappa)}$$

$$= \mathbb{P}(y>0)\, \mathbb{E}[y \mid y>0, \mu, \kappa]$$

(y_2 the underlying NB r.v.) or $= \dfrac{P(y>0)}{1-f(0;\mu,\kappa)} \mathbb{E}[y_2]$ ← doesn't depend on s.

For the ZINB model, which as $f(s;\mu,r)$ is the PMF of a NB distribution with mean μ;

$$\Rightarrow \quad \mathbb{E}[y] = v\mu = q\tau\mu = q\tau e^{\gamma} \quad (\mu = e^{\gamma}).$$

where $\tau = \dfrac{1}{1-(1+\alpha\mu)^{-1/\alpha}} = \dfrac{1}{1-(1+\alpha e^{\gamma})^{-1/\alpha}}$

also $\lim_{r \to -\infty} \tau e^r = \lim_{r \to -\infty} e^r / \tfrac{1}{\tau}$

(l'Hopital)

$$= \lim_{r \to -\infty} \frac{e^r}{\frac{1}{\alpha} \alpha e^r (1 + \alpha e^r)^{-1/\alpha - 1}} \quad \text{(for fixed } \alpha\text{)}$$

$$= \lim_{r \to -\infty} \frac{1}{(1 + \alpha e^r)^{-1/\alpha - 1}} = 1$$

$\Rightarrow \mathbb{E}[Y] \to q$ as $r \to -\infty$

as $r \to \infty$, $\tau \to 1$ for fixed α

$\Rightarrow \mathbb{E}[y] \to q e^r = q\mu$.

On the variance of y in the ZINB.

$$Var(y) = \sum_{s=1}^{\infty} s^2 f(s;\mu,\alpha)\, v - \left(\sum_{r=1}^{\infty} s f(s;\mu,\alpha)\, v \right)^2$$

$$= P(y>0) \sum_{s=1}^{\infty} s^2 \frac{f(s;\mu,\alpha)}{1-f(0;\mu,\alpha)} - \Big(P(y>0)\, \mathbb{E}[y \mid y>0,\mu,\alpha] \Big)^2$$

$$= P(y>0) \Big(var(y \mid y>0,\mu,\alpha) + \mathbb{E}[y \mid y>0,\mu,\alpha]^2 \Big)$$
$$- P(y>0)^2\, \mathbb{E}[y \mid y>0,\mu,\alpha]^2$$

$$= P(y>0)\, var(y \mid y>0,\mu,\alpha) + P(y>0)\big(1-P(y>0)\big)\, \mathbb{E}[y \mid y>0,\mu,\alpha]^2$$

Which in the ZINB case gives, for $f(y;\mu,r)$ the underlying NB, y_2, with $var(y_2;\mu,r) = \frac{\mu}{r}(r+\mu) = \mu + \frac{\mu^2}{r}$, we have,

$\mathbb{E}[y_2^2;\mu,r] = \mu + \mu^2/r + \mu^2$

Since $r = 1/\alpha \Rightarrow \mathbb{E}[y_2^2;\mu,\alpha] = \mu + \mu^2 + \alpha\mu^2$ and $var(y_2;\mu,r) = \mu + \alpha\mu^2$

$$\Rightarrow Var(y) = \underbrace{\frac{P(y>0)}{1-f(0;\mu,\alpha)}}_{=v} \Big\{ \overbrace{var(y_2;\mu,\alpha) + \mathbb{E}[y_2;\mu,\alpha]}^{=\mathbb{E}[y_2^2;\mu,r]} \Big\} - \left(\frac{P(y>0)}{1-f(0;\mu,\alpha)} \right)^2 \mu^2$$

$$= v(\mu + \mu^2 + \alpha\mu^2) - v^2\mu^2$$
$$= v\mu(1 + \mu + \alpha\mu - v\mu)$$

$$\Rightarrow \frac{Var(y)}{\mathbb{E}[y]} = 1 + \mu + \alpha\mu - v\mu$$
$$= 1 + \mu(1 + \alpha - v)$$, assume nontrivial μ

Over dispersion if $1+\alpha-v>0 \Rightarrow -v > -1-\alpha$
$\Rightarrow v < 1+\alpha$; but we know that $v>0$ from its definition (provided $P(y>0) \neq 0$)

$\Rightarrow$ overdispersion if $0 < v < 1+\alpha$

For underdispersion, we have that

$$1 + \mu(1+\alpha - v) < 1 \quad \text{and w/c} \quad var(y) > 0, \; 1+\mu(1+\alpha - v) > 0$$

$$\Rightarrow \quad 1+\alpha - v < 0 \quad \text{and} \quad 1+\alpha - v > -1/\mu$$

$$\Rightarrow \quad -v < -1-\alpha \quad \text{and} \quad -v > -1/\mu - 1 - \alpha$$

$$\Rightarrow \quad v > 1+\alpha \quad \text{and} \quad v < 1/\mu + 1 + \alpha$$

$\Rightarrow$ underdispersion, provided,

$$1+\alpha < v < \frac{1}{\mu} + 1 + \alpha$$

Thus the model is indeed flexible and allows for both under- and over-dispersion.

Finally if $q = 1/\tau = \dfrac{(1+\alpha\mu)^{1/\alpha} - 1}{(1+\alpha\mu)^{1/\alpha}}$ then we recover the NB model.

Proof:

$$g(y;\mu,\alpha) = \begin{cases} 1 - q & y = 0 \\ \dfrac{q}{1-f(0;\mu,\alpha)} \cdot f(y;\mu,\alpha) & y > 0 \end{cases}$$

but $1 - f(0;\mu,\alpha) = \dfrac{(1+\alpha\mu)^{1/\alpha} - 1}{(1+\alpha\mu)^{1/\alpha}} = q$ and $1 - q = 1/(1+\alpha\mu)^{1/\alpha} = f(0;\mu,\alpha)$

$$\Rightarrow g(y;\mu,\alpha) = \begin{cases} f(0;\mu,\alpha) & y = 0 \\ f(y;\mu,\alpha) & y > 0 \end{cases} = f(y;\mu,\alpha)$$

The NB density $\square$.

Implementatn (Following Wood et al. (2017))

According to the framework proposed in Wood et al. (2017), we need the deviance an its derivatives to implement an extend GAM model (see sectn 3.3.1 of that paper);

Differentiating w.r.t. η_i instead of μ_i as this is done in SAI.

1- $\frac{\partial D_i}{\partial \eta_i}, \frac{\partial^2 D_i}{\partial \eta_i^2}, \frac{dh_i}{d\mu_i}, \frac{d^2 h_i}{d\mu_i^2}$ (For $\hat{\vec{\beta}}$)

2- $d^4 h_i/d\eta^4, \frac{\partial^2 D_i}{\partial\theta_i \partial\theta_j}, \frac{\partial^3 D_i}{\partial\eta_i \partial\theta_j \partial\theta_k}, \frac{\partial^4 D_i}{\partial \eta_i^4}, \frac{\partial^4 D_i}{\partial\eta_i^3 \partial\theta_j}, \frac{\partial^4 D}{\partial\eta_i^2 \partial\theta_j \partial\theta_k}$

(For $\hat{\vec{\rho}}$ via full Newton).

As for the ziP model, we first define the deviance as $-2l$ for model estimatn.

w.r.t. η for $y>0$

exactly as for zip.

① $$\frac{dl}{d\eta} = \frac{e^\eta}{e^{e^\eta}-1}, \quad ② \ \frac{d^2 l}{d\eta^2} = (1-e^\eta)\frac{\partial l}{\partial\eta} - \left(\frac{\partial l}{\partial\eta}\right)^2$$

③ $$\frac{d^3 l}{d\eta^3} = -e^\eta \frac{\partial l}{\partial\eta} + (1-e^\eta)^2 \frac{\partial l}{\partial\eta} - 3(1-e^\eta)\left(\frac{\partial l}{\partial\eta}\right)^2 + 2\left(\frac{\partial l}{\partial\eta}\right)^3 \quad \text{and}$$

④ $$\frac{d^4 l}{d\eta^4} = (3e^\eta - 4)e^\eta \frac{\partial l}{\partial\eta} + 4e^\eta\left(\frac{\partial l}{\partial\eta}\right)^2 + (1-e^\eta)^3\frac{\partial l}{\partial\eta} - 7(1-e^\eta)^2\left(\frac{\partial l}{\partial\eta}\right)^2 + 12(1-e^\eta)\left(\frac{\partial l}{\partial\eta}\right)^3 - 6\left(\frac{\partial l}{\partial\eta}\right)^4$$

See SAI of Wood et al. (2016) for dealing with these derivatives as $\eta \to \pm\infty$

w.r.t. γ for $y>0$

① $$\frac{dl}{d\gamma} = \frac{dl^{zt}}{d\gamma} = \frac{d}{d\gamma}\left(y\gamma - y\log(1+\alpha e^\gamma) - \log\{(1+\alpha e^\gamma)^{1/\alpha} - 1\}\right)$$

$$= y - \frac{y}{1+\alpha e^\gamma}\cdot \alpha e^\gamma - \frac{1\cdot \alpha e^\gamma \cdot 1/\alpha\,(1+\alpha e^\gamma)^{1/\alpha - 1}}{(1+\alpha e^\gamma)^{1/\alpha} - 1}$$

$$= y - y\alpha\frac{e^\gamma}{1+\alpha e^\gamma} - \frac{e^\gamma(1+\alpha e^\gamma)^{1/\alpha - 1}}{(1+\alpha e^\gamma)^{1/\alpha} - 1}$$

$$= y - \frac{y\,\alpha e^{\gamma}}{1+\alpha e^{\gamma}} - \left(\frac{e^{\gamma}}{1+\alpha e^{\gamma}}\right)\left(\frac{(1+\alpha e^{\gamma})^{1/\alpha}}{(1+\alpha e)^{1/\alpha} - 1}\right)$$

Define $\beta := \dfrac{e^{\gamma}}{1+\alpha e^{\gamma}}$, $\tau := \dfrac{(1+\alpha e^{\gamma})^{1/\alpha}}{(1+\alpha e^{\gamma})^{1/\alpha} - 1}$

$$\Rightarrow \; d\ell/d\gamma = y - y\alpha\beta - \beta\tau$$

(2) $$d^2\ell/d\gamma^2 = \beta\left\{ y\alpha^2\beta - y\alpha + \alpha\beta\tau - \tau + \beta\tau^2 - \beta\tau \right\}$$

Proof: $$\frac{d}{d\gamma}\left\{ y - y\alpha\beta - \beta\tau \right\} = -y\alpha\, d\beta/d\gamma - \frac{d\beta}{d\gamma}\tau - \beta\frac{d\tau}{d\gamma}$$

But $$d\beta/d\gamma = \frac{(1+\alpha e^{\gamma})e^{\gamma} - e^{\gamma}\cdot\alpha e^{\gamma}}{(1+\alpha e^{\gamma})^2}$$

$$= \frac{e^{\gamma}}{1+\alpha e^{\gamma}} - \frac{\alpha e^{2\gamma}}{(1+\alpha e^{\gamma})^2} = \beta - \alpha\beta^2$$

and $$d\tau/d\gamma = \frac{\left\{\left[(1+\alpha e^{\gamma})^{1/\alpha} - 1\right]\left[\alpha e^{\gamma}\cdot \frac{1}{\alpha}\cdot(1+\alpha e^{\gamma})^{1/\alpha - 1}\right] - \left[(1+\alpha e^{\gamma})^{1/\alpha}\right]\left[\alpha e^{\gamma}\cdot\frac{1}{\alpha}(1+\alpha e^{\gamma})^{1/\alpha-1}\right]\right\}}{\left((1+\alpha e^{\gamma})^{1/\alpha} - 1\right)^2}$$

$$= \frac{e^{\gamma}(1+\alpha e^{\gamma})^{1/\alpha}}{(1+\alpha e^{\gamma})\left[(1+\alpha e^{\gamma})^{1/\alpha} - 1\right]} - \frac{e^{\gamma}(1+\alpha e^{\gamma})^{2/\alpha}}{(1+\alpha e^{\gamma})\left[(1+\alpha e^{\gamma})^{1/\alpha} - 1\right]^2}$$

$$= \beta\tau - \beta\tau^2$$

$$\Rightarrow \frac{d^2\ell}{d\gamma^2} = -y\alpha(\beta - \alpha\beta^2) - \beta\tau + \alpha\beta^2\tau - \beta^2\tau + \beta^2\tau^2$$

$$= y\alpha^2\beta^2 - y\alpha\beta + \alpha\beta^2\tau - \beta\tau + \beta^2\tau^2 - \beta^2\tau$$

$$= \beta\left\{ y\alpha^2\beta - y\alpha + \alpha\beta\tau - \tau + \beta\tau^2 - \beta\tau \right\}$$

$\square$

We also need

$$\mathbb{E}\left[d^2 l / d\gamma^2\right] = \mathbb{E}_Y \, \mathbb{E}\left[d^2 l / d\gamma^2 \,\middle|\, Y\right]$$

$$= \mathbb{E}_Y \left\{ \mathbb{E}\Big[\underbrace{d^2 l / d\gamma^2}_{=0 \text{ for } y=0} \,\Big|\, Y=0\Big] \mathbb{P}(Y=0) + \mathbb{E}\left[\frac{d^2 l}{d\gamma^2} \,\middle|\, Y>0\right] \mathbb{P}(Y>0) \right\}$$

$$= \mathbb{E}_Y \left\{ \mathbb{E}\left[\frac{d^2 l}{d\gamma^2} \,\middle|\, Y>0\right] \underbrace{\mathbb{P}(Y>0)}_{=q} \right\}$$

$$= \mathbb{E}_Y \left\{ q \left(y\alpha^2\beta^2 - y\alpha\beta + \alpha\beta^2\tau - \tau\beta + \beta^2\tau^2 - \beta^2\tau \right) \right\}$$

$$= q \left(\mathbb{E}[y]\alpha^2\beta^2 - \mathbb{E}[y]\alpha\beta + \alpha\beta^2\tau - \tau\beta + \beta^2\tau^2 - \beta^2\tau \right)$$

where $\mathbb{E}[y] = q\tau e^{\gamma}$, as derived on page 6.

(3) $$\frac{d^3 \ell}{d\gamma^3} = \beta \Big\{ -2y\alpha^3\beta^2 + 3y\alpha^2\beta - 2\alpha^2\beta^2\tau - y\alpha + 3\alpha\beta\tau - 3\alpha\beta^2\tau^2 + 3\alpha\beta^2\tau - \tau + 3\beta\tau^2 - 3\beta\tau - 2\beta^2\tau^3 + 3\beta^2\tau^2 - \beta^2\tau \Big\}$$

Proof. Using a numerical linear algebra package, see 👓. $\square$

(4) $$\frac{d^4 \ell}{d\gamma^4} = \beta \Big\{ 6y\alpha^4\beta^3 - 12y\alpha^3\beta^2 + 6\alpha^3\beta^3\tau + 7y\alpha^2\beta - 12\alpha^2\beta^2\tau + 11\alpha^2\beta^3\tau^2 - 11\alpha^2\beta^3\tau - y\alpha + 7\alpha\beta\tau - 18\alpha\beta^2\tau^2 + 18\alpha\beta^2\tau + 12\alpha\beta^3\tau^3 - 18\alpha\beta^3\tau^2 + 6\alpha\beta^3\tau - \tau + 7\beta\tau^2 - 7\beta\tau - 12\beta^2\tau^3 + 18\beta^2\tau^2 - 6\beta^2\tau + 6\beta^3\tau^4 - 12\beta^3\tau^3 + 7\beta^3\tau^2 - \beta^3\tau \Big\}$$

proof. 👓 $\square$

As with $\ell^{(2)}$, some care is required to ensure that the derivatives evaluate accurately without overflow over as wide a range of γ and α as possible. As $\gamma \to \infty$, $\beta = \dfrac{1}{e^{-\gamma} + \alpha} \to 1/\alpha$ for fixed α, while

$$\tau = \frac{1}{1 - (1+\alpha e^{\gamma})^{-1/\alpha}} \to 1, \quad \text{since } 1+\alpha e^{\gamma} \to \infty \Rightarrow (1+\alpha e^{\gamma})^{-1/\alpha} \to 0$$

for fixed α

whereas as $\gamma \to -\infty$, $\beta \to 1/e^{-\gamma} \to 0$ for fixed α, while $e^{\gamma} \to 0$

$\Rightarrow (1+\alpha e^{\gamma})^{1/\alpha} \underset{\substack{\text{first}\\\text{order}\\\text{Taylor}\\\text{expansion}\\\text{about } x=\alpha e^{\gamma}=0}}{\to} 1 + \frac{1}{\alpha}\alpha e^{\gamma} = 1 + e^{\gamma} \Rightarrow \tau = \dfrac{(1+\alpha e^{\gamma})^{1/\alpha}}{(1+\alpha e^{\gamma})^{1/\alpha} - 1}$

$$\to \frac{1+e^{\gamma}}{1+e^{\gamma}-1} = \frac{1+e^{\gamma}}{e^{\gamma}}$$

$$^{*}= e^{-\gamma} + 1 \to e^{-\gamma}$$

* How do we deal with products $\beta\tau$, $\beta^2\tau^2$, $\beta^2\tau$ etc. what does it converge to?

Since as $\gamma \to -\infty$, $\beta \to e^{\gamma}$ and $\tau \to e^{-\gamma}$

$\Rightarrow \beta\tau \to 1$, $\Rightarrow (\beta\tau)^a \to 1 \quad \forall a \in \mathbb{R}$

(alternatively),

$$\lim_{\gamma \to -\infty} \tau_\beta = \lim_{\gamma \to -\infty} \beta \Big/ \left(1 - (1+\alpha e^{\gamma})^{-1/\alpha}\right)$$

$$(\text{l'Hopital}) = \lim_{\gamma \to -\infty} \frac{\beta - \alpha\beta^2}{\frac{1}{\alpha}\alpha e^{\gamma}(1+\alpha e^{\gamma})^{-1/\alpha - 1}}$$

$$= \lim_{\gamma \to -\infty} \frac{\beta(1-\alpha\beta)}{e^{\gamma}(1+\alpha e^{\gamma})^{-1/\alpha - 1}}$$

$$= \lim_{\gamma \to -\infty} \frac{\frac{e^{\gamma}}{1+\alpha e^{\gamma}}(1-\alpha\beta)}{e^{\gamma}(1+\alpha e^{\gamma})^{-1/\alpha - 1}}$$

$$= \lim_{\gamma \to -\infty} \frac{1-\alpha\beta}{(1+\alpha e^{\gamma})^{-1/\alpha}} = 1$$

also; $\beta^a \tau^b \to 0$ if $a > b$ as $\gamma \to -\infty$

as $\gamma \to \infty$, $\beta \to \alpha^{-1}$ and $\tau \to 1$ $\Rightarrow$ $\beta\tau \to \alpha^{-1}$

with these we get

$\partial L/\partial\gamma \to y - 1$ as $\gamma \to -\infty$, $\partial L/\partial\gamma \to y - y - 1/\alpha = -1/\alpha$ as $\gamma \to \infty$

$\partial^2 L/\partial\gamma^2 \to 0$ as $\gamma \to -\infty$, $\partial^2 L/\partial\gamma^2 \to y - y + \alpha/\alpha^2 - \frac{1}{\alpha} + \frac{1}{\alpha^2} - \frac{1}{\alpha^2}$

$= 1/\alpha - 1/\alpha + 1/\alpha^2 - 1/\alpha^2 = 0$

as $\gamma \to \infty$

$\partial^3 L/\partial\gamma^3 \to 0$ as $\gamma \to -\infty$, $\partial^3 L/\partial\gamma^3 \to -2y + 3y - 2/\alpha - y + 3/\alpha$

$- 3/\alpha^2 + 3/\alpha^2 - 1/\alpha + \frac{3}{\alpha^2}$

$- 3/\alpha^2 - 2/\alpha^3$

$+ 3/\alpha^3 - 1/\alpha^3$

$= 1/\alpha - 1/\alpha + 3/\alpha^2 - 3/\alpha^2$

$+ 0 = 0$ as $\gamma \to \infty$

$\partial^4 L/\partial\gamma^4 \to 0$ as $\gamma \to -\infty$, $\partial^4 L/\partial\gamma^4 \to 0$ as $\gamma \to \infty$

For extended GAMs in which η is a Fn of γ and extra parameters θ e.g.,

$$\eta = \theta_1 + (b + e^{\theta_2})\gamma \qquad \text{(as for the ziP model in Wood et al (2016))}$$

$$d\eta/d\theta_1 = 1, \quad d\eta/d\theta_2 = e^{\theta_2}\gamma, \quad d\eta/d\gamma = b + e^{\theta_2}, \quad d^2\eta/d\gamma^2 = 0$$

$$d^2\eta/d\theta_1^2 = d^2\eta/d\theta_1 d\theta_2 = \frac{d^2\eta}{d\gamma\, d\theta_1} = 0, \quad d^2\eta/d\theta_2^2 = e^{\theta_2}\gamma, \quad \frac{d^2\eta}{d\gamma\, d\theta_2} = e^{\theta_2}$$

$$d^3\eta/d\gamma d\theta_1^2 = 0, \quad \frac{d^3\eta}{d\gamma\, d\theta_1\, d\theta_2} = 0, \quad \frac{d^3\eta}{d\gamma d\theta_2^2} = e^{\theta_2},$$

$$d^3\eta/d\gamma^2 d\theta_1 = d^3\eta/d\gamma^2 d\theta_2 = d^3\eta/d\gamma^3 = 0$$

$$d^4\eta/d\gamma^4 = d^4\eta/d\gamma^3 d\theta_1 = d^4\eta/d\gamma^3 d\theta_2 = d^4\eta/d\gamma^2 d\theta_1^2 = \frac{d^4\eta}{d\gamma^2 d\theta_2^2}$$

$$= \frac{d^4\eta}{d\gamma^2 d\theta_1\, d\theta_2} = 0$$

Total derivatives.

I We focus on θ_i, for $i \in \{1,2\}$. derivatives involving θ_0 are considered in sect. 2

$$\frac{d\ell}{d\gamma} = \frac{\partial \ell}{\partial \gamma} + \frac{\partial \ell}{\partial \eta} \cdot \frac{\partial \eta}{\partial \gamma}, \qquad \frac{d^2\ell}{d\gamma^2} = \frac{\partial^2 \ell}{\partial \gamma^2} + \underbrace{\frac{\partial^2 \ell}{\partial \gamma \partial \eta} \cdot \frac{\partial \eta}{\partial \gamma} + \frac{\partial^2 \ell}{\partial \eta \partial \gamma} \frac{\partial \eta}{\partial \gamma}}_{=0} + \frac{\partial^2 \ell}{\partial \eta^2} \cdot \left(\frac{\partial \eta}{\partial \gamma}\right)^2 + \frac{\partial \ell}{\partial \eta} \cdot \frac{\partial^2 \eta}{\partial \gamma^2}$$

$$= \frac{\partial^2 \ell}{\partial \gamma^2} + \frac{\partial^2 \ell}{\partial \eta^2}\left(\frac{\partial \eta}{\partial \gamma}\right)^2 + \frac{\partial \ell}{\partial \eta} \frac{\partial^2 \eta}{\partial \gamma^2}$$

$$\frac{d\ell}{d\theta_i} = \frac{\partial \ell}{\partial \eta} \cdot \frac{\partial \eta}{\partial \theta_i}, \qquad \frac{d^2\ell}{d\gamma\, d\theta_i} = \underbrace{\frac{\partial^2 \ell}{\partial \gamma \partial \eta}}_{=0} \cdot \frac{\partial \eta}{\partial \theta_i} + \frac{\partial \ell}{\partial \eta} \cdot \frac{\partial^2 \eta}{\partial \gamma \partial \theta_i} + \frac{\partial^2 \ell}{\partial \eta^2} \cdot \frac{\partial \eta}{\partial \theta_i} \cdot \frac{\partial \eta}{\partial \gamma}$$

$$= \frac{\partial^2 \ell}{\partial \eta^2} \cdot \frac{\partial \eta}{\partial \theta_i} \cdot \frac{\partial \eta}{\partial \gamma} + \frac{\partial \ell}{\partial \eta} \cdot \frac{\partial^2 \eta}{\partial \gamma \partial \theta_i}$$

$$\frac{d^3\ell}{d\gamma^2 d\theta_i} = \frac{\partial^3 \ell}{\partial \eta^3} \cdot \frac{\partial \eta}{\partial \theta_i} \cdot \left(\frac{\partial \eta}{\partial \gamma}\right)^2 + \frac{\partial^2 \ell}{\partial \eta^2}\, 2 \frac{\partial \eta}{\partial \gamma} \cdot \frac{\partial^2 \eta}{\partial \gamma \partial \theta_i} + \frac{\partial^2 \ell}{\partial \eta^2} \cdot \frac{\partial \eta}{\partial \theta_i} \frac{\partial^2 \eta}{\partial \gamma^2} + \frac{\partial \ell}{\partial \eta} \cdot \frac{\partial^3 \eta}{\partial \gamma^2 \partial \theta_i}$$

$$= \frac{\partial^3 \ell}{\partial \eta^3} \frac{\partial \eta}{\partial \theta_i} \left(\frac{\partial \eta}{\partial \gamma}\right)^2 + \frac{\partial^2 \ell}{\partial \eta^2}\left(2 \frac{\partial \eta}{\partial \gamma} \cdot \frac{\partial^2 \eta}{\partial \gamma \partial \theta_i} + \frac{\partial \eta}{\partial \theta_i} \frac{\partial^2 \eta}{\partial \gamma^2}\right) + \frac{\partial \ell}{\partial \eta} \cdot \frac{\partial^3 \eta}{\partial \gamma^2 \partial \theta_i}$$

$$\frac{d^3\ell}{d\gamma^3} = \frac{\partial^3 \ell}{\partial \gamma^3} + \frac{\partial^3 \ell}{\partial \eta^3} \cdot \frac{\partial \eta}{\partial \gamma} \cdot \left(\frac{\partial \eta}{\partial \gamma}\right)^2 + \frac{\partial^2 \ell}{\partial \eta^2}\left[2\frac{\partial \eta}{\partial \gamma} \cdot \frac{\partial^2 \eta}{\partial \gamma^2}\right]$$

$$+ \frac{\partial^2 \ell}{\partial \eta^2} \cdot \frac{\partial^2 \eta}{\partial \gamma^2} + \frac{\partial \ell}{\partial \eta} \cdot \frac{\partial^3 \eta}{\partial \gamma^3}$$

$$= \partial^3 \ell / \partial \gamma^3 + \frac{\partial^3 \ell}{\partial \eta^3}\left(\frac{\partial \eta}{\partial \gamma}\right)^3 + 3\,\frac{\partial^2 \ell}{\partial \eta^2}\,\frac{\partial^2 \eta}{\partial \gamma^2}\,\frac{\partial \eta}{\partial \gamma} + \frac{\partial \ell}{\partial \eta} \cdot \frac{\partial^3 \eta}{\partial \gamma^3}$$

$$d^2\ell / d\theta_i\, d\theta_j = \frac{\partial^2 \ell}{\partial \eta^2}\ \partial \eta / \partial \theta_j \cdot \partial \eta / \partial \theta_i + \frac{\partial \ell}{\partial \eta} \cdot \partial^2 \eta / \partial \theta_i \partial \theta_j$$

$$\frac{d^3 \ell}{d\gamma\, d\theta_i\, d\theta_j} = \frac{\partial^3 \ell}{\partial \eta^3}\ \partial \eta / \partial \theta_j\ \partial \eta / \partial \theta_i\ \partial \eta / \partial \gamma + \frac{\partial^2 \ell}{\partial \eta^2}\left[\frac{\partial^2 \eta}{\partial \theta_i \partial \theta_j}\,\frac{\partial \eta}{\partial \gamma} + \frac{\partial \eta}{\partial \theta_i}\,\frac{\partial^2 \eta}{\partial \gamma \partial \theta_j}\right]$$

$$+ \frac{\partial^2 \ell}{\partial \eta^2}\,\frac{\partial \eta}{\partial \theta_j}\,\frac{\partial^2 \eta}{\partial \theta_i \partial \gamma} + \frac{\partial \ell}{\partial \eta}\,\frac{\partial^3 \eta}{\partial \gamma \partial \theta_i \partial \theta_j}$$

$$= \partial^3 \ell / \partial \eta^3\ \partial \eta / \partial \theta_i\ \partial \eta / \partial \theta_j\ \partial \eta / \partial \gamma + \partial^2 \ell / \partial \eta^2 \left[\frac{\partial^2 \eta}{\partial \theta_i \partial \theta_j}\,\frac{\partial \eta}{\partial \gamma} + \frac{\partial \eta}{\partial \theta_i}\,\frac{\partial^2 \eta}{\partial \gamma \partial \theta_j}\right.$$

$$\left. + \partial \eta / \partial \theta_j\ \frac{\partial^2 \eta}{\partial \theta_i \partial \gamma}\right] + \partial \ell / \partial \eta\ \partial^3 \eta / \partial \gamma \partial \theta_i \partial \theta_j$$

$$d^4 \ell / d\gamma\, d\gamma\, d\theta_i\, d\theta_j = \frac{\partial^4 \ell}{\partial \eta^4}\,\frac{\partial \eta}{\partial \theta_j}\,\frac{\partial \eta}{\partial \theta_i}\left(\frac{\partial \eta}{\partial \gamma}\right)^2 + \frac{\partial^3 \ell}{\partial \eta^3}\left[\frac{\partial^2 \eta}{\partial \theta_i \partial \theta_j}\left(\frac{\partial \eta}{\partial \gamma}\right)^2 + 2\,\frac{\partial \eta}{\partial \theta_i}\,\frac{\partial \eta}{\partial \gamma}\,\frac{\partial^2 \eta}{\partial \gamma \partial \theta_j}\right]$$

$$+ \frac{\partial^3 \ell}{\partial \eta^3}\,\frac{\partial \eta}{\partial \theta_j}\left(2\,\partial^2 \eta / \partial \gamma \partial \theta_i\ \partial \eta / \partial \gamma + \frac{\partial^2 \eta}{\partial \gamma^2} \cdot \frac{\partial \eta}{\partial \theta_i}\right)$$

$$+ \frac{\partial^2 \ell}{\partial \eta^2}\left(2\,\frac{\partial^3 \eta}{\partial \gamma \partial \theta_i \partial \theta_j}\,\frac{\partial \eta}{\partial \gamma} + \frac{2\partial^2 \eta}{\partial \gamma \partial \theta_i}\,\frac{\partial^2 \eta}{\partial \gamma \partial \theta_j} + \frac{\partial^3 \eta}{\partial \gamma^2 \partial \theta_j}\,\frac{\partial \eta}{\partial \theta_i}\right.$$

$$\left. + \frac{\partial^2 \eta}{\partial \gamma^2}\,\frac{\partial^2 \eta}{\partial \theta_i \partial \theta_j}\right)$$

$$+ \partial^2 \ell / \partial \eta^2\ \partial \eta / \partial \theta_j\ \partial^3 \eta / \partial \gamma^2 \partial \theta_i + \partial \ell / \partial \eta\ \frac{\partial^4 \eta}{\partial \gamma^2 \partial \theta_i \partial \theta_j}$$

$$= \partial^4 l/\partial\eta^4 \; \partial\eta/\partial\theta_j \; \partial\eta/\partial\theta_i \left(\frac{\partial\eta}{\partial\gamma}\right)^2$$

$$+ \partial^3 l/\partial\eta^3 \Big[\partial^2\eta/\partial\theta_i\partial\theta_j \left(\frac{\partial\eta}{\partial\gamma}\right)^2 + 2\, \partial\eta/\partial\theta_i \; \partial\eta/\partial\gamma \; \frac{\partial^2\eta}{\partial\theta_j\partial\gamma}$$

$$+ 2\,\partial\eta/\partial\theta_j \; \partial^2\eta/\partial\gamma\partial\theta_i \; \partial\eta/\partial\gamma + \partial\eta/\partial\theta_i \; \frac{\partial\eta}{\partial\theta_j} \; \frac{\partial^2\eta}{\partial\gamma^2} \Big)$$

$$+ \partial^2 l/\partial\eta^2 \Big[\frac{2\,\partial^3\eta}{\partial\gamma\partial\theta_i\partial\theta_j} \; \partial\eta/\partial\gamma + \frac{2\,\partial^2\eta}{\partial\gamma\partial\theta_i} \; \frac{\partial^2\eta}{\partial\gamma\partial\theta_j} + \frac{\partial^3\eta}{\partial\gamma^2\partial\theta_j} \; \frac{\partial\eta}{\partial\theta_i}$$

$$+ \partial^2\eta/\partial\gamma^2 \; \frac{\partial^2\eta}{\partial\theta_i\partial\theta_j} + \frac{\partial^3\eta}{\partial\gamma^2\partial\theta_i} \; \partial\eta/\partial\theta_j \Big]$$

$$+ \partial l/\partial\eta \; \partial^4\eta/\partial\gamma^2\partial\theta_i\partial\theta_j$$

$$d^4 l/d\gamma^3 d\theta_i = \underbrace{\frac{\partial^4 l}{\partial^3\gamma\partial\eta}}_{=0} \cdot \partial\eta/\partial\theta_i + \frac{\partial^4 l}{\partial\eta^4} \; \frac{\partial\eta}{\partial\theta_i} \left(\frac{\partial\eta}{\partial\gamma}\right)^3 + \frac{\partial^3 l}{\partial\eta^3} \cdot 3\left(\frac{\partial\eta}{\partial\gamma}\right)^2 \frac{\partial^2\eta}{\partial\gamma\partial\theta_i}$$

$$+ \frac{3\,\partial^3 l}{\partial\eta^3} \; \partial\eta/\partial\theta_i \; \partial\eta/\partial\gamma \; \partial^2\eta/\partial\gamma^2 + \frac{3\,\partial^2 l}{\partial\eta^2} \Big[\frac{\partial^2\eta}{\partial\gamma\partial\theta_i} \; \frac{\partial^2\eta}{\partial\gamma^2}$$

$$+ \partial\eta/\partial\gamma \; \frac{\partial^3\eta}{\partial\gamma^2\partial\theta_i} \Big] + \frac{\partial^2 l}{\partial\eta^2} \; \frac{\partial\eta}{\partial\theta_i} \; \frac{\partial^3\eta}{\partial\gamma^3} + \frac{\partial l}{\partial\eta} \; \frac{\partial^4\eta}{\partial\gamma^3\partial\theta_i}$$

$$= \frac{\partial^4 l}{\partial\eta^4} \; \frac{\partial\eta}{\partial\theta_i} \left(\frac{\partial\eta}{\partial\gamma}\right)^3 + 3\,\frac{\partial^3 l}{\partial\eta^3} \Big[\left(\frac{\partial\eta}{\partial\gamma}\right)^2 \frac{\partial^2\eta}{\partial\gamma\partial\theta_i} + \frac{\partial\eta}{\partial\theta_i} \; \frac{\partial\eta}{\partial\gamma} \; \frac{\partial^2\eta}{\partial\gamma^2} \Big]$$

$$+ \partial^2 l/\partial\eta^2 \Big[3\,\partial^2\eta/\partial\gamma\partial\theta_i \; \frac{\partial^2\eta}{\partial\gamma^2} + 3\,\frac{\partial\eta}{\partial\gamma} \; \frac{\partial^3\eta}{\partial\gamma^2\partial\theta_i} + \frac{\partial^3\eta}{\partial\gamma^3} \; \frac{\partial\eta}{\partial\theta_i} \Big]$$

$$+ \partial l/\partial\eta \; \frac{\partial^4\eta}{\partial\gamma^3\partial\theta_i}$$

$$d^4 l/d\gamma^4 = \partial^4 l/\partial\gamma^4 + \frac{\partial^4 l}{\partial\eta^4} \; \partial\eta/\partial\gamma \left(\frac{\partial\eta}{\partial\gamma}\right)^3 + \frac{\partial^3 l}{\partial\eta^3} \cdot 3\left(\frac{\partial\eta}{\partial\gamma}\right)^2 \frac{\partial^2\eta}{\partial\gamma^2}$$

$$+ 3\,\frac{\partial^3 l}{\partial\eta^3} \cdot \partial\eta/\partial\gamma \; \frac{\partial\eta}{\partial\gamma} \; \frac{\partial^2\eta}{\partial\gamma^2} + \frac{3\,\partial^2 l}{\partial\eta^2} \Big[\partial^2\eta/\partial\gamma^2 \; \frac{\partial^2\eta^2}{\partial\gamma^2} +$$

$$\partial\eta/\partial\gamma \; \partial^3\eta/\partial\gamma^3 \Big] + \partial^2 l/\partial\eta^2 \; \partial\eta/\partial\gamma \; \frac{\partial^3\eta}{\partial\gamma^3} + \frac{\partial l}{\partial\eta} \; \frac{\partial^4\eta}{\partial\gamma^4}$$

$$= \frac{\partial^4 \ell}{\partial \gamma^4} + \frac{\partial^4 \ell}{\partial \eta^4}\left(\frac{\partial \eta}{\partial \gamma}\right)^4 + 6\,\frac{\partial^3 \ell}{\partial \eta^3}\left(\frac{\partial \eta}{\partial \gamma}\right)^2 \frac{\partial^2 \eta}{\partial \gamma^2}$$

$$+ \partial^2\ell/\partial\eta^2 \left[3\left(\partial^2\eta/\partial\gamma^2\right)^2 + 4\,\partial\eta/\partial\gamma\ \partial^3\eta/\partial\gamma^3\right) + \frac{\partial \ell}{\partial \eta}\frac{\partial^4 \eta}{\partial \gamma^4}$$

II Derivatives involving θ_0

Since $\alpha = 1/r$ and $r \in \mathbb{R}_{>0} \Rightarrow \alpha \in \mathbb{R}_{>0}$. It is however easier to work with unrestricted parameter. Therefore, we define $\alpha := e^{\theta_0}$ with $\theta_0 \in \mathbb{R}$ unrestricted.

Recall that for $y>0$,

$$\ell^{ZI} = \log\left(1-e^{-e^{\eta}}\right) + y\log\alpha + y\gamma - y\log(1+\alpha e^{\gamma}) - \log\left\{(1+\alpha e^{\gamma})^{1/\alpha}-1\right\}$$
$$+ \log\Gamma(y+1/\alpha) - \log\Gamma(y+1) - \log\Gamma(1/\alpha)$$

$(\alpha = e^{\theta_0})$
$$= \log\left(1-e^{-e^{\eta}}\right) + y\log e^{\theta_0} + y\gamma - y\log(1+e^{\theta_0}e^{\gamma}) - \log\left\{(1+e^{\theta_0}e^{\gamma})^{e^{-\theta_0}}-1\right\}$$
$$+ \log\Gamma(y+e^{-\theta_0}) - \log\Gamma(y+1) - \log\Gamma(e^{-\theta_0})$$

Note that if $x = (1+e^{\theta_0}e^{\gamma})^{e^{-\theta_0}}$, then

$$\log x = e^{-\theta_0}\log(1+e^{\theta_0}e^{\gamma})$$

$$\Rightarrow \frac{1}{x}\,dx/d\theta_0 = \frac{e^{-\theta_0}}{1+e^{\theta_0}e^{\gamma}} \cdot e^{\theta_0}e^{\gamma} + \log(1+e^{\theta_0}e^{\gamma}) \cdot e^{-\theta_0} \cdot -1$$

$$= \frac{e^{\gamma}}{1+e^{\theta_0}e^{\gamma}} - e^{-\theta_0}\log(1+e^{\theta_0}e^{\gamma})$$

$$\Rightarrow dx/d\theta_0 = (1+e^{\theta_0}e^{\gamma})^{e^{-\theta_0}}\left\{\beta - e^{-\theta_0}\log(1+e^{\theta_0}e^{\gamma})\right\}$$

also, $d/dz \log\Gamma(z) =: \psi(z)$, where ψ is the digamma Fn.

Now dropping the ZI superscript, we have,

$$d\ell/d\theta_0 = 0 + y/e^{\theta_0} \cdot e^{\theta_0} + 0 - y/(1+e^{\theta_0}e^{\gamma}) \cdot e^{\theta_0}e^{\gamma} - \frac{1}{(1+e^{\theta_0}e^{\gamma})^{e^{-\theta_0}}-1}$$

$$(1+e^{\theta_0}e^{\gamma})^{e^{-\theta_0}}\left\{\beta - e^{-\theta_0}\log(1+e^{\theta_0}e^{\gamma})\right\}$$

$$-\varphi(y+e^{-\theta_0})e^{-\theta_0} + \varphi(e^{-\theta_0})e^{-\theta_0}$$

Define $\quad \omega := e^{-\theta_0}\log(1+e^{\theta_0}e^{\gamma}) - \beta$

$$= y - y\beta e^{\theta_0} + \tau\omega - e^{-\theta_0}\varphi(y+e^{-\theta_0}) + e^{-\theta_0}\varphi(e^{-\theta_0})$$

To compute $d^2 l/d\theta_0^2$, we need $d\beta/d\theta_0$, and $d\tau/d\theta_0$

$$d\beta/d\theta_0 = \frac{d}{d\theta_0}\frac{e^{\gamma}}{1+e^{\theta_0}e^{\gamma}} = \frac{(1+e^{\theta_0}e^{\gamma})\cdot 0 - e^{\gamma}\cdot e^{\gamma}e^{\theta_0}}{(1+e^{\theta_0}e^{\gamma})^2}$$

$$= -\frac{e^{2\gamma}}{(1+e^{\theta_0}e^{\gamma})^2}e^{\theta_0} = -\beta^2 e^{\theta_0}$$

We leave the rest to the computer; see 👓;

For $\varphi^{(1)}$, see Polygamma Fn on Wikipedia.

$$d^2l/d\theta_0^2 = -y\beta e^{\theta_0} + y\beta^2 e^{2\theta_0} + \tau^2\omega^2 - \tau\omega^2 + \tau(\beta^2 e^{\theta_0} - \omega) + e^{-\theta_0}\varphi(y+e^{-\theta_0}) - e^{-\theta_0}\varphi(e^{-\theta_0}) + e^{-2\theta_0}\varphi^{(1)}(y+e^{-\theta_0}) - e^{-2\theta_0}\varphi^{(1)}(e^{-\theta_0})$$

* How do we deal with products $\tau\omega$ ...

Since as $\gamma \to -\infty$, $\tau \to +\infty$ and $\log(1+e^{\theta_0}e^{\gamma}) \to \log(1) = 0$ we end up with the product $+\infty\cdot 0$, which is indeterminate. But we have hope if $\tau \to +\infty$ slower than the terms it interacts with $\to 0$, in which case the interaction $\to 0$ as $\gamma \to -\infty$.

Taking θ_0 as fixed, we have

(i) $$\lim_{\gamma\to-\infty} \tau\log(1+e^{\theta_0}e^{\gamma}) = \lim_{\gamma\to-\infty}\frac{\log(1+e^{\theta_0}e^{\gamma})}{1-(1+e^{\theta_0}e^{\gamma})^{-e^{-\theta_0}}}$$

(L'Hopital) $$= \lim_{\gamma\to-\infty}\frac{e^{\theta_0}e^{\gamma}}{(1+e^{\theta_0}e^{\gamma})\left(0+e^{-\theta_0}(1+e^{\theta_0}e^{\gamma})^{-e^{-\theta_0}-1}\cdot e^{\theta_0}e^{\gamma}\right)}$$

$$= \lim_{\gamma\to-\infty}\frac{e^{\theta_0}e^{\gamma}}{e^{\gamma}(1+e^{\theta_0}e^{\gamma})^{-e^{-\theta_0}}}$$

$$= \lim_{\gamma \to -\infty} \frac{e^{\theta_0}}{(1+e^{\theta_0}e^{\gamma})^{-e^{-\theta_0}}} = e^{\theta_0}$$

$$\Rightarrow \left(\tau \log(1+e^{\theta_0}e^{\gamma})\right)^a \to e^{a\theta_0} \quad \forall a \in \mathbb{R}$$

also, $\tau^a \log(1+e^{\theta_0}e^{\gamma})^b \to 0$ if $b > a$ as $\gamma \to -\infty$, $a, b \in \mathbb{R}$

(ii) $\lim_{\gamma \to -\infty} \tau\left(e^{-\theta_0}\log(1+e^{\theta_0}e^{\gamma}) - \beta\right) \overset{(i)}{=} e^{-\theta_0} \cdot e^{\theta_0} - 1$ ← $\beta\tau \to 1$, as $\gamma \to -\infty$

$$= 1 - 1 = 0$$

$$\Rightarrow \lim_{\gamma \to -\infty} \tau\omega = 0, \quad \Rightarrow \lim_{\gamma \to -\infty} (\tau\omega)^a = 0 \quad \forall a \in \mathbb{R}$$

$$\Rightarrow \lim_{\gamma \to -\infty} \tau^a \omega^b = 0 \quad \forall a, b \in \mathbb{R} \text{ if } b > a$$

(iii) $\lim_{\gamma \to -\infty} \tau\left(\beta + \beta^2 e^{\theta_0} - e^{-\theta_0}\log(1+e^{\theta_0}e^{\gamma})\right) = \lim_{\gamma \to -\infty} \tau(\beta^2 e^{\theta_0} - \omega)$

$$= \lim_{\gamma \to -\infty} (\beta^2 \tau e^{\theta_0} - \tau\omega) = 0$$

for fixed θ_0

For $\gamma \to +\infty$, $\tau \to 1$, $\beta \to e^{-\theta_0}$, $\log(1+e^{\theta_0}e^{\gamma}) \approx \log(e^{\theta_0}e^{\gamma})$ since $e^{\theta_0}e^{\gamma}$ will dominate 1. and $\log(e^{\theta_0}e^{\gamma}) \to \theta_0 + \gamma$.
We'll use this even though we don't have indeterminacy since the zip implementation also uses it. See how $\log(e^{e^{\gamma}} - 1)$ is approximated as e^{γ} as $\gamma \to \infty$ for the zip model likelihood here 🕶.

With these, we get, as $\gamma \to -\infty$

$$dl/d\theta_0 \to y - e^{-\theta_0}\varphi(y + e^{-\theta_0}) + e^{-\theta_0}\varphi(e^{-\theta_0})$$

$$d^2l/d\theta_0^2 \to e^{-\theta_0}\varphi(y+e^{-\theta_0}) - e^{-\theta_0}\varphi(e^{-\theta_0}) + e^{-2\theta_0}\varphi^{(1)}(y+e^{-\theta_0}) - e^{-2\theta_0}\varphi^{(1)}(e^{-\theta_0})$$

while as $\gamma \to +\infty$

$$\frac{dl}{d\theta_0} \to y - y e^{-\theta_0} e^{\theta_0} - \left(-e^{-\theta_0}(\theta_0+\gamma) + e^{-\theta_0}\right) - e^{-\theta_0}\varphi(y+e^{-\theta_0}) + e^{-\theta_0}\varphi(e^{-\theta_0})$$

$$= e^{-\theta_0}\left\{\theta_0+\gamma - 1 - \varphi(y+e^{-\theta_0}) + \varphi(e^{-\theta_0})\right\}$$

$$\frac{d^2 l}{d\theta_0^2} \to e^{-\theta_0} + e^{-2\theta_0}e^{\theta_0} - e^{-\theta_0}(\theta_0+\gamma) + e^{-\theta_0}\varphi(y+e^{-\theta_0}) - e^{-\theta_0}\varphi(e^{-\theta_0}) + e^{-2\theta_0}\varphi^{(1)}(y+e^{-\theta_0}) - e^{-2\theta_0}\varphi^{(1)}(e^{-\theta_0})$$

$$= e^{-\theta_0}\left\{2 - \theta_0 - \gamma + \varphi(y+e^{-\theta_0}) - \varphi(e^{-\theta_0}) + e^{-\theta_0}\varphi^{(1)}(y+e^{-\theta_0}) - e^{-\theta_0}\varphi^{(1)}(e^{-\theta_0})\right\}$$

$$\frac{d^2 l}{d\theta_0\, d\theta_1} = \frac{d^2 l}{d\theta_0\, d\theta_2} = 0$$

$$\frac{d^2 l}{d\gamma\, d\theta_0} = -y\beta e^{\theta_0} + y\beta^2 e^{2\theta_0} - \beta\tau^2\omega + \beta\tau\omega + e^{\theta_0}\beta^2\tau$$

$$\frac{d^3 l}{d\gamma^2 d\theta_0} = -y\beta e^{\theta_0} + 3y\beta^2 e^{2\theta_0} - 2y\beta^3 e^{3\theta_0} - \beta\tau^2\omega + \beta\tau\omega + 2\beta^2\tau^3\omega + e^{\theta_0}\beta^2\tau^2\omega - 3\beta^2\tau^2\omega - e^{\theta_0}\beta^2\tau\omega + \beta^2\tau\omega + 2e^{\theta_0}\beta^2\tau - 2e^{\theta_0}\beta^3\tau^2 - 2e^{2\theta_0}\beta^3\tau + 2e^{\theta_0}\beta^3\tau$$

$$\frac{d^3 l}{d\gamma\, d\theta_0\, d\theta_1} = \frac{d^3 l}{d\gamma\, d\theta_0\, d\theta_2} = 0$$

$$\frac{d^3 l}{d\gamma\, d\theta_0^2} = -y\beta e^{\theta_0} + 3y\beta^2 e^{2\theta_0} - 2y\beta^2 e^{3\theta_0} - 2\beta\tau^3\omega^2 + 3\beta\tau^2\omega^2 - \beta\tau^2(\beta^2 e^{\theta_0} - \omega) - \beta\tau\omega^2 + \beta\tau(\beta^2 e^{\theta_0} - \omega) + 2\beta^2\tau^2\omega e^{\theta_0} - 2\beta^2\tau\omega e^{\theta_0} + \beta^2\tau e^{\theta_0} - 2\beta^3\tau e^{2\theta_0}$$

$$\frac{d^4 l}{d\gamma^2 d\theta_0\, d\theta_1} = \frac{d^4 l}{d\gamma^2 d\theta_0\, d\theta_2} = 0$$

$$\frac{d^4 l}{d\gamma^2 d\theta_0^2} = -y\beta e^{\theta_0} + 7y\beta^2 e^{2\theta_0} - 12y\beta^3 e^{3\theta_0} + 6y\beta^4 e^{4\theta_0} - 2\beta\tau^3\omega^2 + 3\beta\tau^2\omega^2 - \beta\tau^2(\beta^2 e^{\theta_0} - \omega) - \beta\tau\omega^2 + \beta\tau(\beta^2 e^{\theta_0} - \omega) + 6\beta^2\tau^4\omega^2 + 2\beta^2\tau^3\omega^2 e^{\theta_0} - 12\beta^2\tau^3\omega^2 + 2\beta^2\tau^3(\beta^2 e^{\theta_0} - \omega) - 3\beta^2\tau^2\omega^2 e^{\theta_0} + 7\beta^2\tau^2\omega^2 + 4\beta^2\tau^2\omega e^{\theta_0} + \beta^2\tau^2 e^{\theta_0}(\beta^2 e^{\theta_0} - \omega)$$

(35 terms)

$$-3\beta^2\tau^2\left(\beta^2 e^{\theta_0} - \omega\right) + \beta^2\tau\omega^2 e^{\theta_0} - \beta^2\tau\omega^2 - 4\beta^2\tau\omega e^{\theta_0}$$
$$- \beta^2\tau e^{\theta_0}\left(\beta^2 e^{\theta_0} - \omega\right) + \beta^2\tau\left(\beta^2 e^{\theta_0} - \omega\right) + 2\beta^2\tau e^{\theta_0}$$
$$-8\beta^3\tau^3\omega e^{\theta_0} - 4\beta^3\tau^2\omega e^{2\theta_0} + 12\beta^3\tau^2\omega e^{\theta_0} - 2\beta^3\tau^2 e^{\theta_0}$$
$$+4\beta^3\tau\omega e^{2\theta_0} - 4\beta^3\tau\omega e^{\theta_0} - 8\beta^3\tau e^{2\theta_0} + 2\beta^3\tau e^{\theta_0} + 6\beta^4\tau^2 e^{2\theta_0}$$
$$+6\beta^4\tau e^{3\theta_0} - 6\beta^4\tau e^{2\theta_0}$$

$$\frac{d^4\ell}{d\gamma^3 d\theta_0} = -y\beta e^{\theta_0} + 7y\beta^2 e^{2\theta_0} - 12y\beta^3 e^{3\theta_0} + 6y\beta^4 e^{4\theta_0} - \beta\tau^2\omega + \beta\tau\omega$$
$$+ 6\beta^2\tau^3\omega + 3\beta^2\tau^2\omega e^{\theta_0} - 9\beta^2\tau^2\omega - 3\beta^2\tau\omega e^{\theta_0} + 3\beta^2\tau\omega$$
$$+ 4\beta^2\tau e^{\theta_0} - 6\beta^3\tau^4\omega - 6\beta^3\tau^3\omega e^{\theta_0} + 12\beta^3\tau^3\omega - 2\beta^3\tau^2\omega e^{2\theta_0}$$
$$+ 9\beta^3\tau^2\omega e^{\theta_0} - 7\beta^3\tau^2\omega - 9\beta^3\tau^2 e^{\theta_0} + 2\beta^3\tau\omega e^{2\theta_0} - 3\beta^3\tau\omega e^{\theta_0}$$
$$+ \beta^3\tau\omega - 10\beta^3\tau e^{2\theta_0} + 9\beta^3\tau e^{\theta_0} + 6\beta^4\tau^3 e^{\theta_0} + 9\beta^4\tau^2 e^{2\theta_0}$$
$$- 9\beta^4\tau^2 e^{\theta_0} + 6\beta^4\tau e^{3\theta_0} - 9\beta^4\tau e^{2\theta_0} + 3\beta^4\tau e^{\theta_0}.$$

as $\gamma \to -\infty$ $\quad (\beta \to 0,\ \beta\tau \to 1,\ \tau\omega \to 0)$

$$d^2\ell/d\gamma d\theta_0 \to -0 + 0 - \lim_{\gamma\to-\infty} \beta\tau^2\omega + \underbrace{\lim_{\gamma\to-\infty} \beta\tau\omega}_{=0} + 0$$
$$= \lim_{\gamma\to-\infty}(\beta\tau)\lim_{\gamma\to-\infty}(\tau\omega) = 1 \cdot 0 = 0$$

$$d^3\ell/d\gamma^2 d\theta_0 \to 0, \quad \frac{d^3\ell}{d\gamma d\theta_0^2} \to 0$$

$$d^4\ell/d\gamma^2 d\theta_0^2 \to 0 \text{ and } d^4\ell/d\gamma^3 d\theta_0 \to 0$$

as $\gamma \to +\infty$ $\quad (\beta \to 1/\alpha = e^{-\theta_0},\ \tau \to 1,\ \tilde{x} := \lim_{\gamma\to\infty} x)$

$$\frac{d^2\ell}{d\gamma d\theta_0} \to -y e^{-\theta_0} e^{\theta_0} + y e^{-2\theta_0} e^{2\theta_0} - \widetilde{\beta\omega} + \widetilde{\beta\omega} + e^{\theta_0} e^{-2\theta_0}$$
$$= -y + y - \widetilde{\beta\omega} + \widetilde{\beta\omega} + e^{-\theta_0} = e^{-\theta_0} = 1/\alpha$$

$$d^3\ell/d\gamma^2 d\theta_0 \to -\cancel{y} + \cancel{3}y - \cancel{2}y - \widetilde{\cancel{\beta}\omega} + \widetilde{\cancel{\beta}\omega} + 2\cancel{\beta^2}\omega + e^{\theta_0}\widetilde{\cancel{\beta^2}\omega}$$
$$- 3\widetilde{\cancel{\beta^2}\omega} - \cancel{e^{\theta_0}}\widetilde{\beta^2\omega} + \cancel{\beta^2}\omega + 2e^{\theta_0}\widetilde{\beta^2} - 2\cancel{e^{\theta_0}}\widetilde{\beta^3}$$
$$- 2e^{2\theta_0}\widetilde{\beta^3} + 2\cancel{e^{\theta_0}\widetilde{\beta^3}}$$

$$= 2e^{\theta_0}e^{-2\theta_0} - 2e^{2\theta_0}e^{-3\theta_0}$$
$$= 2e^{-\theta_0} - 2e^{-\theta_0} = 0$$

$$
\frac{d^4\psi_l}{dr^2\,d\theta_0^2} \rightarrow -\cancel{y} + \cancel{7y} - \cancel{12y} + \cancel{6y} - \cancel{2\tilde{\beta}\omega^2} + \cancel{3\tilde{\beta}\omega^2}
$$

$$
- \cancel{\tilde{\beta}^3 e^{\theta_0}} + \cancel{\tilde{\beta}\omega} - \cancel{\tilde{\beta}\omega^2} + \cancel{\tilde{\beta}^3 e^{\theta_0}} - \cancel{\tilde{\beta}\omega}
$$

$$
+ \cancel{6\tilde{\beta}^2\omega^2} + \cancel{2\tilde{\beta}^2\omega^2 e^{\theta_0}} - \cancel{12\tilde{\beta}^2\omega^2} + \cancel{2\tilde{\beta}^4 e^{\theta_0}}
$$

$$
- \cancel{2\tilde{\beta}^2\omega} - \cancel{3\tilde{\beta}^2\omega^2 e^{\theta_0}} + \cancel{7\tilde{\beta}^2\omega^2} + \cancel{4\tilde{\beta}^2\omega e^{\theta_0}}
$$

$$
+ \cancel{\tilde{\beta}^4 e^{2\theta_0}} - \cancel{\tilde{\beta}^2\omega e^{\theta_0}} - \cancel{3\tilde{\beta}^4 e^{\theta_0}} + \cancel{3\tilde{\beta}^2\omega} + \cancel{\tilde{\beta}^2\omega^2 e^{\theta_0}}
$$

$$
- \cancel{\tilde{\beta}^2\omega^2} - \cancel{4\tilde{\beta}^2\omega e^{\theta_0}} - \cancel{\tilde{\beta}^4 e^{2\theta_0}} + \cancel{\tilde{\beta}^2\omega e^{\theta_0}} + \cancel{\tilde{\beta}^4 e^{\theta_0}} - \cancel{\tilde{\beta}^2\omega}
$$

$$
+ 2\tilde{\beta}^2 e^{\theta_0} - \cancel{8\tilde{\beta}^3\omega e^{\theta_0}} - \cancel{4\tilde{\beta}^3\omega e^{2\theta_0}} + \cancel{12\tilde{\beta}^3\omega e^{\theta_0}} - \cancel{2\tilde{\beta}^3 e^{\theta_0}}
$$

$$
+ \cancel{4\tilde{\beta}^3\omega e^{2\theta_0}} - \cancel{4\tilde{\beta}^3\omega e^{\theta_0}} - 8\tilde{\beta}^3 e^{2\theta_0} + \cancel{2\tilde{\beta}^3 e^{\theta_0}} + \cancel{6\tilde{\beta}^4 e^{2\theta_0}}
$$

$$
+ 6\tilde{\beta}^4 e^{3\theta_0} - \cancel{6\tilde{\beta}^4 e^{2\theta_0}}
$$

$$
= 2\tilde{\beta}^2 e^{\theta_0} - 8\tilde{\beta}^3 e^{2\theta_0} + 6\tilde{\beta}^4 e^{3\theta_0}
$$

$$
= 2e^{-\theta_0} - 8e^{-\theta_0} + 6e^{-\theta_0} = 0
$$

$$
\frac{d^4\psi_l}{dr^3\,d\theta_0} \rightarrow -\cancel{y} + \cancel{7y} - \cancel{12y} + \cancel{6y} - \cancel{\tilde{\beta}\omega} + \cancel{\tilde{\beta}\omega} + \cancel{6\tilde{\beta}^2\omega} + \cancel{3\tilde{\beta}^2\omega e^{\theta_0}}
$$

$$
- \cancel{9\tilde{\beta}^2\omega} - \cancel{3\tilde{\beta}^2\omega e^{\theta_0}} + \cancel{3\tilde{\beta}^2\omega} + 4\tilde{\beta}^2 e^{\theta_0} - \cancel{6\tilde{\beta}^3\omega} - \cancel{6\tilde{\beta}^3\omega e^{\theta_0}}
$$

$$
+ \cancel{12\tilde{\beta}^3\omega} - \cancel{2\tilde{\beta}^3\omega e^{2\theta_0}} + \cancel{9\tilde{\beta}^3\omega e^{\theta_0}} - \cancel{7\tilde{\beta}^3\omega} - \cancel{9\tilde{\beta}^3 e^{\theta_0}} + \cancel{2\tilde{\beta}^3\omega e^{2\theta_0}}
$$

$$
- \cancel{3\tilde{\beta}^3\omega e^{\theta_0}} + \cancel{\tilde{\beta}^3\omega} - 10\tilde{\beta}^3 e^{2\theta_0} + \cancel{9\tilde{\beta}^3 e^{\theta_0}} + \cancel{6\tilde{\beta}^4 e^{\theta_0}} + \cancel{9\tilde{\beta}^4 e^{2\theta_0}}
$$

$$
- \cancel{9\tilde{\beta}^4 e^{\theta_0}} + 6\tilde{\beta}^4 e^{3\theta_0} - \cancel{9\tilde{\beta}^4 e^{2\theta_0}} + \cancel{3\tilde{\beta}^4 e^{\theta_0}}
$$

$$
= 4\tilde{\beta}^2 e^{\theta_0} - 10\tilde{\beta}^3 e^{2\theta_0} + 6\tilde{\beta}^4 e^{3\theta_0}
$$

$$
= 4e^{-\theta_0} - 10e^{-\theta_0} + 6e^{-\theta_0} = 0
$$

On sampling.

For a zero-truncated density, f^*;

$$f^*(x) = \begin{cases} \dfrac{f(x)}{1-F(0)} & \text{for } x > 0, \\ 0 & \text{otherwise} \end{cases}$$

where $F(0) = \mathbb{P}(X \leq 0)$, in our case (discrete pmf)

$F(0) = \mathbb{P}(X=0) + \mathbb{P}(X<0)$
$= f(0) + 0 = f(0)$

the truncated CDF is

$$F^*(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ \displaystyle\int_0^x \frac{f(u)}{1-F(0)}\,du & \text{for } x > 0 \end{cases}$$

$$= \begin{cases} 0 & \text{for } x \leq 0 \\ \dfrac{F(x) - F(0)}{1 - F(0)} & \text{for } x > 0 \end{cases}$$

To sample from this distribution, we proceed as follows;

(P0) 1. Generate $U \sim \mathcal{U}(0,1)$
2. Let $V = F(0) + (1-F(0))\,U$
3. Return $X = F^{-1}(V)$

Proof. We show that $X \sim F^*$

$$\mathbb{P}(X \leq x) = \mathbb{P}(F^{-1}(V) \leq x) = \mathbb{P}(V \leq F(x))$$
$$= \mathbb{P}\big(F(0) + (1-F(0))\,U \leq F(x)\big)$$
$$= \mathbb{P}\left(U \leq \frac{F(x) - F(0)}{1 - F(0)}\right)$$
$$= \begin{cases} 0 & \text{if } x < 0 \quad (\text{since } F(x) = 0 \text{ and } \mathbb{P}(U \leq -x) = 0) \\ \dfrac{F(x)-F(0)}{1-F(0)} & \text{for } x > 0 \end{cases}$$
$$= F^*(x) \quad \square$$

(Since $U \sim \mathcal{U}(0,1)$
$\Rightarrow \mathbb{P}(U \leq u) = u$)

$V \sim \mathcal{U}(F(0), 1)$

proof.
$$\mathbb{P}(V \leq v) = \mathbb{P}\big(F(0) + (1-F(0))\,U \leq v\big)$$
$$= \mathbb{P}\left(U \leq \frac{v - F(0)}{1-F(0)}\right)$$
$$= \begin{cases} 0 & \text{if } v < F(0) \\ \dfrac{v - F(0)}{1-F(0)} & \text{if } F(0) \leq v \leq 1 \\ 1 & \text{if } v > 1 \end{cases}$$
$$= \mathcal{U}(F(0), 1) \qquad \square$$

So, the procedure above can be reduced to
(P1)

1 Generate $V \sim \mathcal{U}(F(0), 1)$
2 Return $X = F^{-1}(V)$

again $\mathbb{P}(X \leq x) = \mathbb{P}(F^{-1}(V) \leq x) = \mathbb{P}(V \leq F(x))$
$$= \begin{cases} 0 & \text{if } F(x) \leq F(0) \Leftrightarrow x \leq 0 \\ \dfrac{F(x) - F(0)}{1 - F(0)} & \text{if } F(x) > F(0) \Leftrightarrow x > 0 \end{cases}$$
$$= F^*(x)$$

For the ZT negative binomial.
Recall the ZT NB density

$$f(y \mid y > 0, r, p) = \frac{\Gamma(y+r)}{\Gamma(y+1)\Gamma(r)} \, p^r (1-p)^y \Big/ 1 - p^r$$

Some care is need for sampling.

1. as $r \to 0$, $f(y|y>0, r, p)$ is not defined b/c $1-p^r \to 0$

Note, $\frac{\Gamma(y+r)}{\Gamma(y+1)\Gamma(r)} = \frac{\Gamma(y+r)}{y!\,\Gamma(r)}$

and using Euler's reflection formula. 🕶️

$\Gamma(r)\Gamma(1-r) = \pi / \sin \pi r$, as $r \to 0$ $\Gamma(1-r) \to \Gamma(1) = 0! = 1$
and $\sin \pi r \to \pi r$

$\Rightarrow \Gamma(r) \to 1/r$

$\Rightarrow \frac{\Gamma(y+r)}{\Gamma(y+1)\Gamma(r)} \to \frac{\Gamma(y)}{y!\ 1/r} = \frac{r\,(y-1)!}{y!} = \frac{r}{y}.$

as the first order Maclaurin expansion of p^r around $r=0$, is

$p^r \approx p^0 + r \frac{dp^r}{dr}\Big|_{r=0} = 1 + r\left(p^r \ln p \big|_{r=0}\right)$
$= 1 + r \ln p$

let $y = p^r$
$\Rightarrow \ln y = r \ln p$
$\Rightarrow \frac{1}{y} \frac{dy}{dr} = \ln p$
$\Rightarrow dy/dr = y \ln p$
$= p^r \ln p$

$\Rightarrow 1 - p^r \xrightarrow{r \to 0} -r \ln p$

$\Rightarrow f(y|y>0, r, p) \xrightarrow[r \to 0]{} \frac{r}{y}(1-p)^y / -r\ln p = -\frac{(1-p)^y}{y \ln p}$

This is exactly the logarithmic distribution with parameter $(1-p)$

So for small p, i.e., as $p \to 0$, sample from $\mathrm{Log}(1-p)$

2. $p \to 1$,

again $f(y \mid y>0, r, p)$ is not defined;

Note that as $p \to 1$

$$P(y=1 \mid y>0, r, p) = f(1 \mid y>0, r, p)$$

$$= \frac{\Gamma(1+r)}{\Gamma(1+1)\,\Gamma(r)} \, \frac{p^r (1-p)^1}{1-p^r}$$

$$= \frac{r!}{(r-1)!} \, \frac{p^r (1-p)}{1-p^r} = \frac{r p^r (1-p)}{1-p^r} \qquad 1-p^r \simeq r(1-p)$$

Note that $1-p^r$ first order Taylor expanded about $p=1$ is

$$1-p^r \approx (1-1^r) + \left.\frac{d(1-p^r)}{dp}\right|_{p=1} (p-1) = 0 + -r p^{r-1}\Big|_{p=1} p-1$$

$$= -r(p-1) = r(1-p)$$

$$\Rightarrow \frac{r p^r (1-p)}{1-p^r} \approx \frac{r p^r (1-p)}{r(1-p)} = p^r \to 1 \text{ as } p \to 1$$

$$\Rightarrow p(y=1 \mid y>0, r, p) \to 1 \text{ as } p \to 1.$$

* Why do we take care of these cases?

One may wonder, since the truncated distribution F^* isn't used directly in the sampling procedure. So why did we derive these limits ($r \to 0$, $p \to 1$). It is b/c as $r \to 0$ or $p \to 1$, $F(0)$ ($= f(0; r, p)$), in our case) goes to 1 (p^0 or 1^r). Thus the support of V would be very small ($[F(0), 1]$). Thus the samples from $U(F(0), 1)$ would be fairly indistinguishable (and close to 1) due to machine precision and the quantile $F^{-1}(V)$ may generate values that are not 1 due to machine errors, but we would want 1s in these cases. Secondly, by resolving these limits we can directly set the value of our samples (e.g for $p \to 1$) or sample directly from a known distribution (for $r \to 0$), which is more efficient than the procedure (lower runtime).

Sampling from hurdle model.

The hurdle is of the form

$$g(y) = \begin{cases} 1-p & \text{for } y = 0 \\ \dfrac{p\, f(y)}{1-F(0)} & \text{for } y > 0 \end{cases} \qquad \xrightarrow[\text{CDF}]{\text{simple } \int_0^y g(y)\,dy} \qquad G(y) = \begin{cases} 1-p, & y=0 \\ p\,\dfrac{F(y)-F(0)}{1-F(0)}, & y>0 \end{cases}$$

To sample from this, we follow the following procedure.

(P2), see Grimmett ch. 4.11 Example 7.

1. Generate $u \sim U(0,1)$ and compute p.
2. If $u \le 1-p$ set $Y \leftarrow 0$
3. otherwise generate $X \sim G(\cdot, u) = F^{*}$, by setting $X = F^{-1}(V)$, V being as in (P1) and set $Y \leftarrow X$

Proof. We prove that this procedure generates samples from the hurdle model.

$$P(Y \le 0) = P(Y=0) = P(U \le 1-p) = 1-p \quad \text{-- } 1$$

and for $y \ge 0$

$$P(Y \le y) = P(U > 1-p) \cdot P(X \le y; U > 1-p)$$

(u here is different from u used in (P0))

2nd NB.

$$= 1 - P(U \le 1-p)\; F^{*}(y), \quad \leftarrow \text{since } y \ge 0 \text{ (see def of } F^{*})$$

$$= p\,F^{*}(y) = p\left(\frac{F(y)-F(0)}{1-F(0)}\right) \text{ -- } 2$$

. 1 and 2 gives $G(y)$ $\square$

For use in `predict` (se-fit)

Recall $\mathbb{E}[Y] = q\,\tau\,e^{\gamma} = q\,\dfrac{e^{\gamma}}{1-(1+\alpha e^{\gamma})^{-1/\alpha}}$; $Y \sim$ ZTNB.

Let the ZTNB mean be $\bar{\mu}$,

$$\bar{\mu} = \frac{e^{\gamma}}{1-(1+\alpha e^{\gamma})^{-1/\alpha}}, \quad d\bar{\mu}/d\gamma = e^{\gamma}\left[\frac{d\tau}{d\gamma}\right] + \tau e^{\gamma}$$

$$= e^{\gamma}(\beta\tau - \beta\tau^2) + \tau e^{\gamma}$$

$$= e^{\gamma}\beta\tau - e^{\gamma}\beta\tau^2 + \tau e^{\gamma}$$

as $\gamma \to -\infty$

$$e^{\gamma}\tau \to 1$$

$$e^{\gamma}\beta\tau \to \beta, \quad e^{\gamma}\beta\tau^2 \to e^{\gamma}\tau \text{ (since } \beta\tau \to 1)$$

$$\to 1 \quad (\text{as } \tau \to e^{-\gamma})$$

$$\Rightarrow \bar{\mu} \to \beta - 1 + 1 = \beta$$

as $\gamma \to \infty$, $\bar{\mu} = (e^{\gamma}\beta\tau - e^{\gamma}\beta\tau^2 + \tau e^{\gamma})$

$$\to e^{\gamma}/\alpha - e^{\gamma}/\alpha + e^{\gamma}$$

$$= e^{\gamma}$$

$$q = 1 - e^{-e^{\eta}}, \quad dq/d\gamma = dq/d\eta \cdot d\eta/d\gamma$$

$$dq/d\eta = e^{-e^{\eta}}e^{\eta}, \quad \eta = \theta_1 + (b + e^{\theta_2})\gamma$$

$$\Rightarrow d\eta/d\gamma = b + e^{\theta_2}$$

$$\Rightarrow dq/d\gamma = e^{-e^{\eta}}e^{\eta}(b + e^{\theta_2})$$, (we don't do asymptotics with η b/c it is endogenous i.e. it depends on γ already)

$$\Rightarrow \frac{d\mathbb{E}[Y]}{d\gamma} = dq/d\gamma \cdot \bar{\mu} + q \cdot \frac{d\bar{\mu}}{d\gamma}$$