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

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

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

$\square$

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

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

$\Rightarrow \mu p = r - rp$

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

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

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

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

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

is well defined

Define $r := 1/\alpha$,

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

$\Rightarrow$ The single data log-likelihood is

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

Zero-truncated NB.

The probability of a zero count is

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

The zero-truncated density is

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

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

This gives us the single data log-likelihood;

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

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. y a erated

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

where $q$ is the probability of a positive counts

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

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

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

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

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

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

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

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 $\log(1-e^{-e^{\eta}}) \to \log(e^{\eta} - e^{2\eta}/_2 + e^{3\eta}/_6) \to \eta$ 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 $\alpha$ as fixed.

Proof:
① $\log(1-e^{-e^{\eta}}) \to \log\left(e^{\eta} - e^{2\eta}/_2 + e^{3\eta}/_6\right) \to \eta$ $n \to -\infty$
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 $x \to 0$

$$\Rightarrow e^{-e^{\eta}} \to 1 - e^{\eta} + e^{2\eta}/_2 - e^{3\eta}/_6$$

$\Rightarrow$ 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 $e^{\eta}$

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

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

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

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

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

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

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

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

as $\gamma \to -\infty$

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

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

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 $r \in \mathbb{N}_0$

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

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

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

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

## On the expectation of $y$ in the ZINB.

Recall the model's density

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

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

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

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

Therefore, the mean is given by

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

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

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

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

($y_2$ the underlying NB r.v.)

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

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

## On the variance of $y$ in the ZINB.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$\Rightarrow$ 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:

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

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

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

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. $\eta_i$ instead of $\mu_i$ as this is done in SA I.

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

(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. $\eta$ for $y>0$

Exactly as for zip.

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

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

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

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

w.r.t. $\gamma$ for $y>0$

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

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

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

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

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

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

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

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

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

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

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

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

$\square$

We also need

$$\mathbb{E}[d^2l/d\gamma^2] = \mathbb{E}_Y\, \mathbb{E}[d^2l/d\gamma^2 \mid Y]$$

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

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

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

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

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

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

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

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

**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 $\gamma$ and $\alpha$ as possible. As $\gamma \to \infty$, $\beta = \frac{1}{e^{-\gamma}+\alpha} \to 1/\alpha$ for fixed $\alpha$, while

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

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

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

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

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

* 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),

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

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

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

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

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

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 $\gamma \to -\infty$, $\partial^2 l / \partial \gamma^2 \to y - y + \alpha/\alpha^2 - 1/\alpha + 1/\alpha^2 - 1/\alpha^2$

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

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 $\gamma \to \infty$

$\partial^4 l / \partial \gamma^3 \to 0$ as $\gamma \to -\infty$, $\partial^4 l / \partial \gamma^4 \to \cancel{-2y} + \cancel{3y} - \cancel{2/\alpha} - \cancel{y} + \cancel{3/\alpha}$

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

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

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

For extended GAMs in which $\eta$ is a Fn of $\gamma$ and extra parameters $\theta$ e.g.;

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

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

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

In this setting

$$dl/d\gamma = \partial l/\partial\gamma + \partial l/\partial\eta \cdot \partial\eta/\partial\gamma, \quad \frac{d^2 l}{d\gamma^2} = \frac{\partial^2 l}{\partial\gamma^2} \overbrace{+ \frac{\partial^2 l}{\partial\gamma\partial\eta} \cdot \frac{\partial\eta}{\partial\gamma}}^{=0}$$

$$\left. + \frac{\partial^2 l}{\partial\eta\partial\gamma} \frac{\partial\eta}{\partial\gamma} \right\} = 0$$

$$+ \frac{\partial^2 l}{\partial\eta^2} \cdot \left(\frac{\partial\eta}{\partial\gamma}\right)^2$$

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

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

$$dl/d\theta_i = \partial l/\partial\eta \cdot \partial\eta/d\theta_i, \quad \frac{d^2 l}{d\gamma d\theta_i} = \underbrace{\frac{\partial^2 l}{\partial\gamma\partial\eta}}_{=0} \cdot \frac{\partial\eta}{d\theta_i} + \frac{\partial l}{\partial\eta} \cdot \frac{\partial^2\eta}{\partial\gamma\partial\theta_i} + \frac{\partial^2 l}{\partial\eta^2} \cdot \frac{\partial\eta}{\partial\theta_i} \cdot \partial\eta/\partial\gamma$$

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

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

$$+ \partial l/\partial\eta \cdot \partial^3\eta/\partial\gamma^2\partial\theta_i$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

## Derivatives w.r.t. $\theta_0$.

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 $y > 0$,

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

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

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

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

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

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

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

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

Now dropping the ZI superscript, we have,

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

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

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

To compute $d^2l/d\theta_0^2$, 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 👓;

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

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

* How do we deal with products $\tau e^{-\theta}\log(1+e^{\theta_0}e^{\gamma})$, ...
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 of interacts with $\to 0$, in which case the interaction $\to 0$ as $\gamma \to -\infty$.
Taking $\theta_0$ 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) $$= \lim_{\gamma\to-\infty}\frac{e^{\theta_0}e^{\gamma}}{(1+e^{\theta_0}e^{\gamma})\left(0+e^{-\theta_0}(1+e^{\theta_0}e^{\gamma})^{-e^{-\theta_0}-1}\cdot e^{\theta_0}e^{\gamma}\right)}$$

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

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

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

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

$\leftarrow \beta\tau \to 1$, as $\gamma \to -\infty$

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

$$= 1 - 1 = 0$$

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

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

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

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

$$= 1 + 0 - e^{-\theta_0}e^{\theta_0} = 1 - 1 = 0$$

for fixed $\theta_0$

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 implementatn also uses it. See how $\log(e^{e^{\gamma}}-1)$ is approximated as $e^{\gamma}$ 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 $\gamma \to +\infty$

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

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

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

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