Annex A.
Theoretical results to the paper: "Some findings on zero-inflated and hurdle Poisson models for disease mapping"

Francisca Corpas-Burgos
Fundación para el Fomento de la Investigación Sanitaria y Biomedica de la Comunidad Valenciana (FISABIO), Valencia, Spain

Gonzalo García-Donato
Departamento de Análisis Económico y Finanzas, Universidad de Castilla la Mancha, Albacete, Spain

Miguel A. Martinez-Beneito
Fundación para el Fomento de la Investigación Sanitaria y Biomedica de la Comunidad Valenciana (FISABIO), Valencia, Spain and CIBER de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain

This document shows four different general results on the modeling of zero inflated and hurdle Poisson models with either fixed or random effects. At the end we draw two corollaries with some specific results of particular interest for the paper above.

Result 1. *Let* $\mathbf{O} = \{O_i : i = 1, ..., I\}$ *be independent observations from the hurdle Poisson model*

$$O_i \sim (1 - \pi_i(\boldsymbol{u}, \sigma))^{1_{\{0\}}(O_i)} \left(\pi_i(\boldsymbol{u}, \sigma) \frac{Poi(O_i \mid E_i R_i)}{1 - Poi(0 \mid E_i R_i)}\right)^{1_{(0,\infty)}(O_i)},$$

where

$$\pi_i(\boldsymbol{u}, \sigma) = F(\sigma \boldsymbol{z}_i \boldsymbol{u}), \;\; with \;\; \boldsymbol{u} \sim f(\boldsymbol{u}) = N_I(\boldsymbol{0}, \boldsymbol{I}),$$

being F *a distribution function with* $F(-x) = 1 - F(x)$ *and* $\{\boldsymbol{z}_i : i = 1, ..., I\}$ *a set of* I*-dimensional vectors. Let also* $\boldsymbol{Z}^\star$ *be the* $I \times I$ *matrix with rows* $\boldsymbol{z}_i^\star$ *defined as* $\boldsymbol{z}_i$ *if* $O_i = 0$ *or* $-\boldsymbol{z}_i$ *if* $O_i > 0$*. Assume that* σ*,* $\boldsymbol{u}$ *and* $\boldsymbol{R}$ *are independent a priori and* σ *follows an improper prior distribution* $f(\sigma)$*. Let*

$$\mathcal{C} = \{\boldsymbol{v} \in \mathcal{R}^I : \; \boldsymbol{Z}^\star \boldsymbol{v} \leq 0\},$$

if the following condition is satisfied

$$dimension\,(\mathcal{C}) = I, \tag{1}$$

then the posterior distribution $f(\boldsymbol{u}, \sigma, \boldsymbol{R} \mid \boldsymbol{O})$ *is improper independently on the prior distribution* $f(\boldsymbol{R})$ *assumed for* $\boldsymbol{R}$*.*

Proof. The proof uses a similar technique as the proof for impropriety of posterior distributions in Bernoulli experiments derived in [1] (Theorem 1.i).

We have to show that the integral

$$\int L(\boldsymbol{u}, \sigma, \boldsymbol{R}; \boldsymbol{O})\, f(\boldsymbol{u}, \sigma, \boldsymbol{R})\, d\boldsymbol{u} d\sigma d\boldsymbol{R}$$

diverges, where L is the likelihood function, that is,

$$L(\boldsymbol{u}, \sigma, \boldsymbol{R}; \boldsymbol{O}) = \prod_{i=1}^{I} (1 - \pi_i(\boldsymbol{u}, \sigma))^{1_{\{0\}}(O_i)} \left(\pi_i(\boldsymbol{u}, \sigma) \frac{Poi(O_i \mid E_i R_i)}{1 - Poi(0 \mid E_i R_i)} \right)^{1_{(0,\infty)}(O_i)}$$

and f is the prior distribution. It can be easily seen that the integral above is:

$$\int \prod_{\{i:O_i=0\}} \frac{Poi(O_i \mid E_i R_i)}{1 - Poi(0 \mid E_i R_i)} \Big\{ \int \int_{\mathcal{R}^I} \prod_{\{i:O_i=0\}} (1 - F(\sigma \boldsymbol{z}_i \boldsymbol{u})) \prod_{\{i:O_i>0\}} F(\sigma \boldsymbol{z}_i \boldsymbol{u}) f(\sigma) f(\boldsymbol{u}) d\boldsymbol{u} d\sigma \Big\} f(\boldsymbol{R})\, d\boldsymbol{R}.$$

Bearing in mind that $F(-x) = 1 - F(x)$ the inner integral above results:

$$\int \int_{\mathcal{R}^I} \prod_{\{i:O_i=0\}} (1 - F(\sigma \boldsymbol{z}_i^* \boldsymbol{u})) \prod_{\{i:O_i>0\}} F(-\sigma \boldsymbol{z}_i^* \boldsymbol{u}) f(\sigma) f(\boldsymbol{u}) d\boldsymbol{u} d\sigma =$$

$$\int \int_{\mathcal{R}^I} \prod_{i=1}^{I} (1 - F(\sigma \boldsymbol{z}_i^* \boldsymbol{u})) f(\sigma) f(\boldsymbol{u}) d\boldsymbol{u} d\sigma \geq \int \int_{\mathcal{C}} \prod_{i=1}^{I} (1 - F(\sigma \boldsymbol{z}_i^* \boldsymbol{u})) f(\sigma) f(\boldsymbol{u}) d\boldsymbol{u} d\sigma \geq$$

$$\geq \int f(\sigma) d\sigma \int_{\mathcal{C}} \frac{1}{2^I} f(\boldsymbol{u}) d\boldsymbol{u} = \frac{1}{2^I} \int f(\sigma) d\sigma.$$

The last integral obviously diverges if $f(\sigma)$ is improper.

□

Result 2. *Let* $\mathbf{O} = \{O_i : i = 1, ..., I\}$ *be independent observations from the hurdle Poisson model*

$$O_i \sim (1 - \pi_i(\boldsymbol{\beta}))^{1_{\{0\}}(O_i)} \left(\pi_i(\boldsymbol{\beta}) \frac{Poi(O_i \mid E_i R_i)}{1 - Poi(0 \mid E_i R_i)} \right)^{1_{(0,\infty)}(O_i)},$$

where

$$\pi_i(\boldsymbol{\beta}) = F(\boldsymbol{x}_i \boldsymbol{\beta}),$$

$\boldsymbol{x}_i = (x_{i1}, \ldots, x_{iJ})$ *are* J*-dimensional vectors of known covariates and* F *a distribution function. Suppose that* $\boldsymbol{\beta}$ *is, a priori, independent of* $\boldsymbol{R}$ *with prior distribution*

$$\boldsymbol{\beta} \sim \prod_{j=1}^{J} f_j(\beta_j), \ \ \beta_j \in \mathcal{R}.$$

If for any $1 \leq j^* \leq J$, $x_{ij^*} > 0$ *for all* i *with* $O_i > 0$ *and negative otherwise (respectively* $x_{ij^*} > 0$ *for all* i *with* $O_i = 0$ *and negative otherwise) and* $\int f(\beta_{j^*}) d\beta_{j^*}$ *diverges for large positive (respectively negative) values of* β_{j^*} *then the posterior distribution* $f(\boldsymbol{\beta}, \boldsymbol{R} \mid \boldsymbol{O})$ *is improper independently on the prior distribution* $f(\boldsymbol{R})$ *assumed for* $\boldsymbol{R}$*.*

Proof. Let us assume the case $x_{ij^*} > 0$ for all i with $O_i > 0$ and negative otherwise. The likelihood function can be put as

$$L(\boldsymbol{\beta}, \boldsymbol{R}; \boldsymbol{O}) = \prod_{\{i:O_i=0\}} (1-\pi_i(\boldsymbol{\beta})) \prod_{\{i:O_i>0\}} \pi_i(\boldsymbol{\beta}) \frac{Poi(O_i \mid E_i R_i)}{1-Poi(0 \mid E_i R_i)} =$$

$$\left(\prod_{\{i:O_i=0\}} (1-F(\boldsymbol{x}_i\boldsymbol{\beta})) \prod_{\{i:O_i>0\}} F(\boldsymbol{x}_i\boldsymbol{\beta}) \right) \prod_{\{i:O_i>0\}} \frac{Poi(O_i \mid E_i R_i)}{1-Poi(0 \mid E_i R_i)}.$$

Thus,

$$\int_R L(\boldsymbol{\beta}, \boldsymbol{R}; \boldsymbol{O}) f(\beta_{j^*}) d\beta_{j^*} > \int_0^\infty L(\boldsymbol{\beta}, \boldsymbol{R}; \boldsymbol{O}) f(\beta_{j^*}) d\beta_{j^*} \propto$$

$$\int_0^\infty \prod_{\{i:O_i=0\}} (1-F(\boldsymbol{x}_i\boldsymbol{\beta})) \prod_{\{i:O_i>0\}} F(\boldsymbol{x}_i\boldsymbol{\beta}) f(\beta_{j^*}) d\beta_{j^*}.$$

Since F is a distribution function is also, in particular, an increasing function. Moreover, as $x_{ij^*} > 0$ for all i with $O_i > 0$ and negative otherwise we have that, if $\boldsymbol{\beta}_0 = (\beta_1, ..., \beta_{j^*-1}, 0, \beta_{j^*+1}, ..., \beta_J)$, this expression is greater than

$$\prod_{\{i:O_i=0\}} \left(1-F(\boldsymbol{x}_i\boldsymbol{\beta}_0')\right) \prod_{\{i:O_i>0\}} F(\boldsymbol{x}_i\boldsymbol{\beta}_0') \int_0^\infty f(\beta_{j^*}) d\beta_{j^*}$$

which diverges due to the prior impropriety of $f(\beta_{j^*})$ for large positive values.

The proof for the case $x_{ij^*} > 0$ if $O_i = 0$ and negative otherwise is analogous. □

Result 3. *Let* $\mathbf{O} = \{O_i : i = 1, ..., I\}$ *be independent observations from the ZIP model*

$$O_i \sim (1-\pi_i(\boldsymbol{u}, \sigma)) 1_{\{0\}}(O_i) + \pi_i(\boldsymbol{u}, \sigma)\, Poi(O_i \mid E_i R_i),$$

where

$$\pi_i(\boldsymbol{u}, \sigma) = F(\sigma \boldsymbol{z}_i^t \boldsymbol{u}), \;\; \boldsymbol{u} \sim f(\boldsymbol{u}) = N_I(\boldsymbol{0}, \boldsymbol{I}_I),$$

being F *a distribution function with* $F(-x) = 1 - F(x)$ *and* $\{\boldsymbol{z}_i, i = 1, ..., I\}$ *a set of* I*-dimensional vectors. Let also* $\boldsymbol{Z}^\star$ *be the* $I \times I$ *matrix with rows* $\boldsymbol{z}_i^\star$ *defined as* $\boldsymbol{z}_i$ *if* $O_i = 0$ *or* $-\boldsymbol{z}_i$ *if* $O_i > 0$*. Assume that* σ, $\boldsymbol{u}$ *and* $\boldsymbol{R}$ *are independent a priori and* σ *follows an improper prior distribution* $f(\sigma)$*. Let*

$$\mathcal{C} = \{\boldsymbol{v} \in \mathcal{R}^I : \; \boldsymbol{Z}^\star \boldsymbol{v} \leq 0\},$$

if the following condition is satisfied

$$dimension\,(\mathcal{C}) = I,$$

then the posterior distribution $f(\boldsymbol{u}, \sigma, \boldsymbol{R} \mid \boldsymbol{O})$ *is improper independently on the prior distribution* $f(\boldsymbol{R})$ *assumed for* $\boldsymbol{R}$*.*

Proof. The Likelihood function for this model can be expressed as:

$$L(\boldsymbol{u},\sigma,\boldsymbol{R};\boldsymbol{O}) = \prod_{i=1}^{I}(1-\pi_i(\boldsymbol{u},\sigma)+\exp(-E_iR_i))^{1_{\{0\}}(O_i)}(\pi_i(\boldsymbol{u},\sigma)Poi(O_i|E_iR_i))^{1_{(0,\infty)}(O_i)} \geq$$

$$\prod_{i=1}^{I}(1-\pi_i(\boldsymbol{u},\sigma))^{1_{\{0\}}(O_i)}(\pi_i(\boldsymbol{u},\sigma)Poi(O_i|E_iR_i))^{1_{(0,\infty)}(O_i)},$$

which is proportional, as a function of $\boldsymbol{u}$ and σ to the likelihood function of the hurdle Poisson model. Since the conditions of this results are the same than for Result 1 and there $\int f(\boldsymbol{u},\sigma,\boldsymbol{R} \mid \boldsymbol{O})d\boldsymbol{u}d\sigma$ diverged, it follows that $f(\boldsymbol{u},\sigma,\boldsymbol{R} \mid \boldsymbol{O})$ is now improper as a direct consequence of that Result. □

Result 4. *Let* $\mathbf{O} = \{O_i : i = 1, ..., I\}$ *be independent observations from the ZIP model*

$$O_i \sim (1-\pi_i(\boldsymbol{\beta}))1_{\{0\}}(O_i) + \pi_i(\boldsymbol{\beta})\,Poi(O_i \mid E_iR_i),$$

where

$$\pi_i(\boldsymbol{\beta}) = F(\boldsymbol{x}_i\boldsymbol{\beta}),$$

$\boldsymbol{x}_i = (x_{i1},\ldots,x_{iJ})$ *are J-dimensional vectors of known covariates and F a distribution function. Suppose that* $\boldsymbol{\beta}$ *is, a priori, independent of* $\boldsymbol{R}$ *with prior distribution*

$$\boldsymbol{\beta} \sim \prod_{j=1}^{J} f_j(\beta_j), \ \ \beta_j \in \mathcal{R}.$$

If for any $1 \leq j^* \leq J$, $x_{ij^*} > 0$ *for all* $i = 1,\ldots,I$ *(respectively* $x_{ij^*} < 0$*) and* $\int f(\beta_{j^*})d\beta_{j^*}$ *diverges for large positive (respectively negative) values of* β_{j^*} *then* $f(\boldsymbol{\beta},\boldsymbol{R} \mid \boldsymbol{O})$ *is improper independently on the prior distribution* $f(\boldsymbol{R})$ *assumed for* $\boldsymbol{R}$*.*

Proof. Let us assume the case $x_{ij^*} > 0$ for all i. The likelihood function is

$$L(\boldsymbol{\beta},\boldsymbol{R};\boldsymbol{O}) = \prod_{i=1}^{I}\big((1-\pi_i(\boldsymbol{\beta}))1_{\{0\}}(O_i) + \pi_i(\boldsymbol{\beta})\,Poi(O_i \mid E_iR_i)\big).$$

The above expression corresponds to a sum with 2^I positive terms. Obviously if the integral of any one of these terms times the prior is divergent then the posterior distribution would be improper. One of these terms in the likelihood is

$$L_1(\boldsymbol{\beta},\boldsymbol{R};\boldsymbol{O}) = \prod_{i=1}^{I} Poi(O_i \mid E_iR_i)\pi_i(\boldsymbol{\beta}) = \prod_{i=1}^{I} Poi(O_i \mid E_iR_i)F(\boldsymbol{x}_i\boldsymbol{\beta}).$$

The function $F(\cdot)$ is increasing and therefore, as $x_{ij^*} > 0$, is also an increasing function of β_{j^*}. Then, if $\boldsymbol{\beta}_0 = (\beta_1,\ldots,\beta_{j^*-1},0,\beta_{j^*+1},\ldots,\beta_J)$

$$L_1(\boldsymbol{\beta},\boldsymbol{R};\boldsymbol{O}) \geq L_1(\boldsymbol{\beta}_0,\boldsymbol{R};\boldsymbol{O})$$

for any $\boldsymbol{\beta}$ with $\beta_{j^*} > 0$. Then

$$\int L_1(\boldsymbol{\beta}, \boldsymbol{R}; \boldsymbol{O}) f(\boldsymbol{\beta})\, d\beta_{j^*} \geq L_1(\boldsymbol{\beta}_0, \boldsymbol{R}; \boldsymbol{O}) \int_0^\infty f_{j^*}(\beta_{j^*})\, d\beta_{j^*},$$

and, since $f_{j^*}(\beta_{j^*})$ diverges for large positive values, $f(\boldsymbol{\beta}, \boldsymbol{R} \mid \boldsymbol{O})$ is improper.

The proof of the Result for $x_{ij} < 0$ is analogous. □

Corollary 1. *Let us consider a hurdle Poisson model with Poisson means modeled as a BYM model and probabilities of zeroes as*

$$logit(\pi_i) = \boldsymbol{x}_i \boldsymbol{\beta} + v_i,$$

for $\boldsymbol{v} \sim N_I(\boldsymbol{0}, \sigma^2 \boldsymbol{C})$ a vector of random effects with $\boldsymbol{C}$ a symmetric, positive-definite structure matrix and $\boldsymbol{x}_i$ a vector of covariates of length J. Then,

1. *If $f(\sigma)$ is improper then $f(\boldsymbol{\beta}, \sigma, \boldsymbol{u}, \boldsymbol{R}|\boldsymbol{O})$ diverges regardless of $f(\beta, \boldsymbol{u}, \boldsymbol{R})$.*

2. *Let us assume $\boldsymbol{\beta}$ to be a priori independent with $\beta_j \sim f(\beta_j)$ $j = 1, ..., J$, and there is a j^* $(1 \leq j^* \leq J)$ with $x_{ij^*} > 0$ when $O_i > 0$ and negative otherwise (respectively $x_{ij^*} > 0$ when $O_i = 0$ and negative otherwise). If $\int f(\beta_{j^*}) d\beta_{j^*}$ diverges for large positive (respectively negative) values, then $f(\boldsymbol{\beta}, \sigma, \boldsymbol{u}, \boldsymbol{R}|\boldsymbol{O})$ is improper regardless of $f(\sigma, \boldsymbol{u}, \boldsymbol{R})$.*

3. *Both previous results also hold for:*
   - *probit or tobit link functions for modelling $\boldsymbol{\pi}$.*
   - *non-Poisson discrete likelihoods (such as binomial or negative-binomial).*
   - *other spatial structures, beyond BYM, for the Poisson means.*

Proof. .

Proof for item 1:

This is just a particular case of Result 1 for $F(x) = antilogit(x)$, which is the distribution function for a logistic density. Moreover, the linear term for $\boldsymbol{\pi}$ is a bit different since we have now random effects v_i instead of $\sigma \boldsymbol{z}_i \boldsymbol{u}$ and fixed effects.

Regarding $\boldsymbol{v}$, if $\boldsymbol{C} = \boldsymbol{\Lambda} \boldsymbol{D}^2 \boldsymbol{\Lambda}'$ is the eigendecomposition of $\boldsymbol{C}$, then $\boldsymbol{v} = \sigma \boldsymbol{Z} \boldsymbol{u}$ for $\boldsymbol{u} \sim N_I(\boldsymbol{0}_I, \boldsymbol{I}_I)$ and $\boldsymbol{Z} = \boldsymbol{\Lambda} \boldsymbol{D}$. In this case the matrix $\boldsymbol{Z}^*$ in Result 1 would be $\boldsymbol{Z}^* = \boldsymbol{L}\boldsymbol{Z}$ for $\boldsymbol{L}$ a diagonal matrix with $L_{ii} \in \{-1, 1\}$ for all i. Since $\boldsymbol{C}$ is positive definite then $\boldsymbol{Z}$ is of full rank. Besides, since both $\boldsymbol{L}$ and $\boldsymbol{Z}$ are full rank, then $\boldsymbol{Z}^*$ is also full rank and therefore regular so $\boldsymbol{y}\boldsymbol{Z}^* = \boldsymbol{0}$ iif $\boldsymbol{y} = \boldsymbol{0}$. [1] (Section 2.4) state that Condition (1) holds iif there is no nonnegative vector, $\boldsymbol{y} \neq \boldsymbol{0}$ such that $\boldsymbol{y}\boldsymbol{Z}^* = \boldsymbol{0}$ so, according to this criterion, that condition holds also for this Corollary.

Finally, regarding the fixed effects in the linear term for $\boldsymbol{\pi}$. The last integral in Result 1 would be now of the form

$$\int\int_{\mathcal{R}^I}\prod_{i=1}^{I}(1-F(\boldsymbol{x}_i\boldsymbol{\beta}+\sigma\boldsymbol{z}_i^*\boldsymbol{u}))f(\sigma)f(\boldsymbol{u})d\boldsymbol{u}d\sigma$$

which by similar bounding arguments as those used in the last part of the proof of Result 1, is a divergent integral.

Proof for item 2:
This is just a particular case of Result 2 for $F(x) = antilogit(x)$ and with an additional random effects term in the linear predictor. This term would not interfere at all in last integral of the proof of that result which makes the posterior distribution improper. So this result keeps being valid with the additional random effects term.

Proof for item 3:
First note that using probit or tobit link functionswould be equivalent to consider normal or t probability density functions for $F()$. So these would be also particular cases of Result 1. Note also that the Poisson likelihood does not have any effect on the posterior impropriety of the proofs of Results 1 and 2 so this could also be changed to binomial or negative-binomial distributions, for example. Finally, note that the BYM model for the Poisson means is irrelevant for the posterior impropriety in these models since the impropriety comes from their zero-specific terms. □

Corollary 2. *Let us consider a ZIP model with Poisson means modeled as a BYM model and probabilities of extra-Poisson zeroes as*

$$logit(\pi_i) = \boldsymbol{x}_i\boldsymbol{\beta} + v_i,$$

for $\boldsymbol{v} \sim N_I(\boldsymbol{0}, \sigma^2\boldsymbol{C})$ *a vector of random effects with* $\boldsymbol{C}$ *a symmetric, positive-definite full-rank structure matrix and* $\boldsymbol{x}_i$ *a vector of covariates of length* J*. Then,*

1. *If* $f(\sigma)$ *is improper then* $f(\boldsymbol{\beta}, \sigma, \boldsymbol{u}, \boldsymbol{R}|\boldsymbol{O})$ *diverges regardless of* $f(\beta, \boldsymbol{u}, \boldsymbol{R})$.

2. *Let us assume* $\boldsymbol{\beta}$ *to be a priori independent with* $\beta_j \sim f(\beta_j)$ $j = 1, ..., J$*, and there is a* j^* $(1 \leq j^* \leq J)$ *with* $x_{ij^*} > 0$ *for* $i = 1, ..., I$ *(respectively* $x_{ij^*} < 0$ *for* $i = 1, ..., I$*). If* $\int f(\beta_{j^*})d\beta_{j^*}$ *diverges for large positive (respectively negative) values then* $f(\boldsymbol{\beta}, \sigma, \boldsymbol{u}, \boldsymbol{R}|O)$ *is improper regardless of* $f(\sigma, \boldsymbol{u}, \boldsymbol{R})$.

3. *Both previous results also hold for:*
   - *probit or tobit link functions for modelling* $\boldsymbol{\pi}$.
   - *non-Poisson discrete likelihoods (such as binomial or negative-binomial).*
   - *other spatial structures beyond BYM for the Poisson means.*

Proof. .
Follow the same argument than for Corollary 1 applied to Results 3 and 4.

□

References

[1] Ranjini Natarajan and Charles E. McCulloch. A note on the existence of the posterior distribution for a class of mixed models for binomial responses. *Biometrika*, 82(3):639–643, 1995.