Supplementary Material

# Box S1: Conway-Maxwell-Poisson Regression

A two-parameter generalization of the Poisson distribution was introduced by Conway and Maxwell (1962). This distribution, now known as the COM-Poisson distribution, can model data that are either under- or over-dispersed relative to the Poisson distribution (Sellers and Shmueli, 2010; Shmueli et al., 2005). It is characterized by two parameters: **** is the location parameter and **** is the dispersion parameter that outlines deviations from Poisson distribution. Value of **ν** corresponds to under-dispersion if ****1 and over-dispersion if ****1; ****1 yields the Poisson distribution.

In the original formulation, the probability mass function for the COM-Poisson is given by the equation (S1):

|  |  |
| --- | --- |
|  | (S1) |

Where Pr(Y = y| ,) is the conditional probability mass function of the event Y = y, given the parameters  and ; *Z*(,) is a normalizing factor.

Mean and variance are estimated by (S2):

|  |  |
| --- | --- |
|  | (S2) |

where *E*[Y] is the expectation (or mean) for the event Y and *VAR*[Y] is its variance.

A limitation of the COM-Poisson model is that neither **** nor **** represent a clear centering parameter. To overcome this drawback Guikema & Goffelt (2008) proposed a parametrization of the original COM-Poisson model, using the relation (S3):

|  |  |
| --- | --- |
|  | (S3) |

With this parametrization, **µ** is a reasonable approximation of the mean, while **** keeps its role as a shape parameter.

Using this parametrization, a maximum-likelihood estimation method was used for fitting the COM-Poisson under a generalized linear model (GLM) framework, as follows (S4):

|  |  |
| --- | --- |
|  | (S4) |

Where ***E*[Y|X]** is the expectation of the count random variable Y, given the set X of independent variables xi; **β0** and **βi** are regression parameters,**** is the centering parameter and **** is the shape parameter, which can be a constant or a function.

With this formulation, the shape parameter **ν** can be open to a second link function to allow the amount of over-dispersion or under-dispersion to vary across measurements.

For large values of variable Y or for ν close to 1, equation (S4) may be reduced to equation (S5)

|  |  |
| --- | --- |
|  | (S5) |

For N observations, the log-likelihood for the GLM model can be written as (S6):

|  |  |
| --- | --- |
|  | (S6) |

Where L(Y|X, β,ν) is the log-likelihood of the set Y, given X, , ν; L(Y|X, β, ν) is the objective-function that should be maximized during GLM regression; yi is the observed value, i is the location parameter calculated from function (S3) and  is the shape parameter.

Note that the likelihood function of the COM-Poisson regression model, differently from the Poisson model, includes the **ν** parameter. It follows that the log-likelihood functions could be quite different even when the β parameters are similar.

# References

Conway, R.W., & Maxwell, W.L., (1962). A queueing model with state dependent service rate. Journal of Industrial Engineering, 12, 132-136.

Guikema, S.D., & Goffelt, J.P. (2008). A flexible count data regression model for risk analysis. Risk Analysis, 28, 213-223.

Sellers, K.F., & Shmueli, G. (2010). A flexible regression model for count data. Annals of Applied Statistics, 4, 943-961. https://doi.org/10.1214/09-AOAS306.

Shmueli, G., Minka, T., Kadane, J.B., Borle, S., & Boatwright, P.B. (2005). A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution. Journal of the Royal Statistical Society, Series C, 54, 127-42.

# Box S2: text Equations (8) and (9)

Equation (8), which is a second-degree polynomial equation, can be solved for ***E***[**Nt]** and **v** (wxMaxima, Version 17.10.1, <http://wxmaxima.sourceforge.net/>), but solutions (S7 and S8) are not very comfortable:

|  |  |
| --- | --- |
|  | (S7) |

and

|  |  |
| --- | --- |
|  | (S8) |

However, under certain circumstances, Equation (8) can be simplified. When **ν** approaches to 1, the term **1/(2 ν) -1/2** become null and Equation (8) is reduced to Equation (9). When the term **1/(2 ν) -1/2** corresponds to the maximum acceptable absolute error **e*E* [Nt],** we can write (S9):

|  |  |
| --- | --- |
|  | (S9) |

Where **e** is the maximum relative acceptable error.

Then:

|  |  |
| --- | --- |
|  | (S10) |

|  |  |
| --- | --- |
|  | (S11) |
|  | (S12) |

replacing **ν** in Equation (9) we have (S13):

|  |  |
| --- | --- |
|  | (S13) |

|  |  |
| --- | --- |
|  | (S14) |

|  |  |
| --- | --- |
|  | (S15) |

From (S15) it is evident that **c0 = 2 e**. Then, when **c0 ≤ 2 e**, Equation (8) can be simplified to Equation (9), as in the text.

# Box S3: Empirical approach for selecting the additional non-randomness model component

In response to the osmotic treatment a transitory additional variation, not explained by the equation 7 and possibly attributable to aggregation, occurred over time. This additional variation was higher at the beginning of the process and then decreased along with the decreasing of the number of survivors. Therefore, the appearance of this variation and its behaviour during treatment required the introduction of an additional model component, defined as “additional non-randomness” variation. To describe this component, a number of empirical equations were developed (Table 1-Box S3) and the Akaike information criterion (AIC) was used to select the best fit model under parsimony. AIC was calculated using the number of the fitted parameters and the number of components included in the additional non-randomness variation. A single model, which gave the best fit to the data, was chosen.

**Table 1-Box S3. Different models and their AIC**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **c0** | **εt** | **t^** | ***E*[Nt]^** | **1/*E*[N0]^** | **LL** | **n** | **p** | **AIC** |  |
| 0.0181 | 0.0128 | 1 | 2 | 0.5 | -2854.8 | 993 | 5 | 5719.6 | \* |
| 0.0315 | 0.0181 | 1 | 1 | -1 | -2833.7 | 993 | 5 | 5677.4 | \*\* |
| 0.0234 | 0.0290 | 1 | 1.5 | -- | -2845.6 | 993 | 4 | 5699.2 | \*\* |
| 0.0346 | 0.0487 | 1 | 1 | -- | -2848.4 | 993 | 4 | 5704.7 | \*\* |
| 0.0183 | 0.0060 | 1 | 2 | -- | -2855.4 | 993 | 4 | 5718.7 | \*\* |
| 0.0190 | 0.0255 | 1 | 2 | 1 | -2855.2 | 993 | 5 | 5720.4 | \*\* |
| -- | -- | -- | 1 | 0.5 | -- | 993 | 4 | -- | \*\*\* |
| 0.0269 | 0.0648 | 1 | 2 | 2 | -2860.6 | 993 | 5 | 5731.2 | \*\*\*\* |
| 0.0339 | 0.0044 | 2 | 2 | 2 | -2864.0 | 993 | 5 | 5737.9 | \*\*\*\* |
| 0.0405 | 0.0616 | 1 | 1 | 1 | -2865.0 | 993 | 5 | 5739.9 | \*\*\*\* |
| 0.0455$ | -- | -- | -- | -- | -2870.7 | 993 | 1 | 5743.4 | \*\*\*\* |

\*selected model (equation 12)

\*\*models excluded for variance to mean ratio unrealistic prediction

\*\*\*models excluded for lack of convergence

\*\*\*\* models excluded for AIC performance

n: numbers of data

p: numbers of parameters and components