FaBE 2013 International Conferences on Food and Biosystems Engineering

(SET) (CD-ROM) ISBN: 978-960-9510-09-7

Proceedings of FaBE 2013 International Conferences on Food and Biosystems Engineering

Editors : Dr. K. Petrotos and Dr. A. Filintas

Skiathos, 30 May - 02 June 2013, GREECE.

[Volume 1]

[vol. 1] ISBN: 978-960-9510-10-3

Technological Educational Institute of Thessaly
Greece. **[vol. 1] ISBN: 978-960-9510-10-3**

[eng]

Proceedings of FaBE 2013 International Conferences on Food and Biosystems Engineering

Editors : Dr. K. Petrotos and Dr. A. Filintas

Skiathos Island, GREECE

30 May - 02 June 2013.

[Volume 1]

[vol. 1] ISBN: 978-960-9510-10-3

Published by the **Technological Educational Institute of Thessaly (TEI of Thessaly), [Greece].**

All papers of the present volume were peer reviewed by two independent reviewers. Acceptance was granted when both reviewers' recommendations were positive.

FaBE 2013 International Conferences on

Food and Biosystems Engineering

(SET) (CD-ROM) ISBN: 978-960-9510-09-7

AN ALTERNATIVE FRAMEWORK TO CONDUCT INFERENTIAL STATISTICS FOR LOW MICROBIAL COUNTS IN FOODS: THE POISSON-GAMMA REGRESSION

Ursula Gonzales-Barron[1], Vasco Cadavez[1*] and Francis Butler[2]
[1] CIMO Mountain Research Centre, Polytechnic Institute of Braganza, Braganza, Portugal
[2] UCD School of Biosystems Engineering, University College Dublin, UCD Belfield, Ireland
* Corresponding author: vcadavez@ipb.pt

ABSTRACT

The objective of this article was to compare four Poisson-gamma regression models to assess the effect of chilling on the concentration of coliforms from beef carcasses. A total of 600 carcasses were sampled before and after chilling at eight large Irish abattoirs, and the total coliforms were determined. With a coded variable (pre-chill/post-chill) as treatment, and extracting the variability of batches nested in abattoirs, random-effects models confirmed that chilling had a decreasing effect on the overall recovery of coliforms. Furthermore, the expected coliforms concentrations on pre-chill and post-chill carcasses were estimated on a CFU/cm^2 scale, as well as their between-batch variability. This study introduced an alternative conceptual framework that can find interesting applications in stochastic risk assessment and in the design of more efficient sampling plans.

Keywords: Negative binomial, chilling, coliforms, carcass.

1. INTRODUCTION

In the evaluation of microbiological quality of foodstuffs, bacterial concentration is conventionally expressed in terms of logarithm of the colony forming units (CFU). Logarithmic transformation is believed to approximate data normality, which is fundamental for the application of inferential statistical data analysis such as the analysis of variance (ANOVA). This may lead to the widely-held practice that whenever bacterial colonies are not observed (zero counts), a low log value corresponding to the limit of enumeration of the microbiological test can be used. This statistical practice for 'censored' observations is known as imputation, and, depending on the proportion of zero counts or censored points, the mean values are normally overestimated. On the other hand, Gonzales-Barron and Butler (2011a,b) have demonstrated that the lognormal (and Poisson-lognormal) distributions are only appropriate for the representation of high microbial counts, while the Poisson-gamma distribution performs much better for low microbial counts. Thus, for bacterial counts of lower occurrence (such as some hygiene indicators and pathogens), an alternative conceptual framework based on the Poisson-gamma distribution should be explored.

The Poisson-gamma models correspond to a family of flexible count data regressions that can model more variability than the nominal variance under the simple Poisson model (which is equal to the expected value). Heterogeneity or clustering causes a condition called over-dispersion, meaning that the variance of the observed count data normally exceeds the mean. A heterogeneous Poisson model loosens the Poisson restriction by allowing the expected number of counts (λ) to be a function of some random variable. If this random variable follows a gamma distribution, the resulting heterogeneous Poisson will be a Poisson-gamma, also known as a negative binomial distribution. The objective of this work was to introduce an alternative count data regression framework to conduct inferential statistics on microbial counts that did not approximate to a normal distribution after logarithmic transformation. Specifically, the effect of chilling on the coliforms counts recovered from Irish beef carcasses, was assessed by fitting mixed-effects Poisson-gamma regression models.

2. METHODOLOGY

2.1 Sampling of pre-chill and post-chill beef carcasses and microbiological analysis

Eight beef export Irish abattoirs were visited, and a total of 600 beef carcasses were sampled for the quantification of the total coliforms present at the pre-chill and post-chill stage (Gonzales-Barron et al., 2010). All abattoirs were visited at least twice, and during each visit, 30 animals were randomly sampled at the end of the slaughter line.

2.2 Poisson-gamma regression model

In its simplest form, the Poisson distribution models the number of events from a memory-less exponential process where the event rate λ is constant. Assuming that (i) there are no losses in the transfer of bacterial cells from the swabs to the homogenate; (ii) the bacterial cells extracted from the swabs are randomly distributed in the 200-ml neat homogenate; and (iii) each of the plated cells will become a colony after incubation, let Y_i be the random variable for the number of bacterial colonies *i*, counted on a Petri dish. If Y_i follows a Poisson distribution, the probability mass function is,

$$\Pr(Y_i) = \frac{Exp(-\mu_i) \times \mu_i}{Y_i!} \qquad \ldots(1)$$

$$\mu_i = \frac{A}{V} dt \times \lambda_i \qquad \ldots(2)$$

where A is the swabbed area of the carcass (cm^2), which was measured for each of the sampled carcasses, V is the homogenate volume (200 ml), d is the dilution level at which the respective plate count Y_i was made (i.e., 10^0, 10^{-1}, 10^{-2}, etc.), t represents the aliquot volume poured onto the Petri dish (1 ml), and λ_i is the unknown mean bacterial concentration in cell/cm^2.

In a Poisson regression model, the mean parameter λ_i would be a function of a vector of covariates X_n, where β is a parameter vector. In our case, the only covariate is a coded variable X (0 as pre-chill, 1 as post-chill) to quantify the effect of the chilling treatment β_1.

$$\lambda_i = \exp(\beta_0 + \beta_1 X) \qquad \ldots(3)$$

However, carcasses within a batch do not share the same true unknown microbial concentration λ_i (CFU/cm^2) due to the heterogeneity in contamination. Thus, the basic Poisson regression model is generalised by including a dispersion parameter to accommodate the heterogeneity in the count data. Then, a generalised Poisson distribution lets the expected microbial concentration λ_i be a function also of some unobserved random variable e_i (Hinde and Demetrio, 1998),

$$\lambda_i = \exp(\beta_0 + \beta_1 X + e_i) = \exp(\beta_0 + \beta_1 X)\exp(e_i) \qquad \ldots(4)$$

When $exp(e_i)$ follows a gamma distribution $\Gamma(1/k,\ k)$ with expected value 1 and dispersion parameter k, the above modification to the Poisson distribution yields a more flexible distribution, the Poisson-gamma. The discrete probability mass function for the microbial plate count Y can then be estimated as,

$$\Pr(Y_i) = \frac{\Gamma(Y_i + k^{-1})}{\Gamma(Y_i + 1)\Gamma(k^{-1})} \left[\frac{k^{-1}}{k^{-1} + \lambda_i}\right]^{k^{-1}} \left[\frac{\lambda_i}{k^{-1} + \lambda_i}\right]^{Y_i} \qquad \ldots(5)$$

where $\lambda_i = \exp(\beta_0 + \beta_1 X)$ and the gamma function $\Gamma(a) = \int_0^\infty e^{-s} s^{a-1} ds$. The Poisson-gamma and Poisson models are considered to be nested because as *k* approaches to zero, the Poisson-gamma converges to Poisson.

2.3 Fitted variants of Poisson-regression models

2.3.1 Model 1

This model presumes that the measure of dispersion *k* is common for all the sampled batches of beef carcasses. Nevertheless, this model takes into account that there is some heterogeneity in the observed plate counts *Y* that originate from the production batches *b* nested in the different abattoirs *a*. To extract this source of heterogeneity, the fitted function of the expected microbial concentration λ becomes,

$$\lambda = \exp(\beta_0 + \lfloor \beta_1 + u_{b(a)} \rfloor X) \qquad \ldots(6)$$

where the chilling effect β_1 is allowed to take up different independent values batch to batch. The nested random effects $u_{b(a)}$ are assumed to have a normal distribution with mean zero and standard deviation σ_u.

2.3.2 Model 2

This model does not only extract the batch-to-batch variability in the chilling effect β1 but also in the intercept β0. Thus, the expected microbial concentration becomes,

$$\lambda = \exp(\lfloor \beta_0 + v_{b(a)} \rfloor + \lfloor \beta_1 + u_{b(a)} \rfloor X) \qquad \textbf{...(7)}$$

where the random effects $v_{b(a)}$ model the shifts in the intercept β_0 for the logarithm of the contamination level in each batch. The $v_{b(a)}$ are assumed to have a normal distribution with mean zero and standard deviation σ_v. The correlation ρ between both random effects was assessed.

2.3.3 Model 3

This is a flexible model in which both the mean λ and the dispersion parameter k are affected by chilling. The model then becomes,

$$\lambda = \exp\left(\beta_0 + \left\lfloor \beta_1 + u_{b(a)} \right\rfloor X\right)$$
$$k = \exp\left(\alpha_0 + \alpha_1 X\right) \qquad ...(8)$$

where the intercept of the logarithmic expression defining k is λ_0, and the slope describing the chilling effect on k is α_1.

2.3.4 Model 4

In this model, the between-batch heterogeneity in the chilling effect affects both the mean λ and the dispersion parameter k. The model then becomes,

$$\lambda = \exp\left(\beta_0 + \left\lfloor \beta_1 + u_{b(a)} \right\rfloor X\right)$$
$$k = \exp\left(\alpha_0 + \left\lfloor \alpha_1 + v_{b(a)} \right\rfloor X\right) \qquad ...(9)$$

The nested random-effects $v_{b(a}$ are assumed to have a normal distribution with mean zero and standard deviation σ_v. The correlation ρ between both random effects was assessed.

The four models were fitted to the coliforms counts data set using the SAS procedure PROC NLMIXED for non-linear mixed models (SAS version 9.1, SAS Institute Inc, NC, USA).

3. RESULTS AND DISCUSSION

As forecast by Gill et al. (1998), who pointed out that for there to be an approximation to normal distribution, the bacteria of interest should be counted in at least 85% of the samples, the microbial data set of coliforms in beef, consisting of 67% of positive counts, did not fit to a normal distribution. For plotting the histogram of frequencies in log CFU/cm^2, the zero counts were replaced by the microbiological limit of enumeration of −2.0 log CFU/cm^2. As shown in Figure 1, neither for the pre-chill nor the post-chill group, the microbial concentration could be approximated to a normal distribution. Here, it is worthy to mention that such practice of replacing zero counts by the limit of enumeration has been very common for statistical treatment either merely descriptive or inferential.

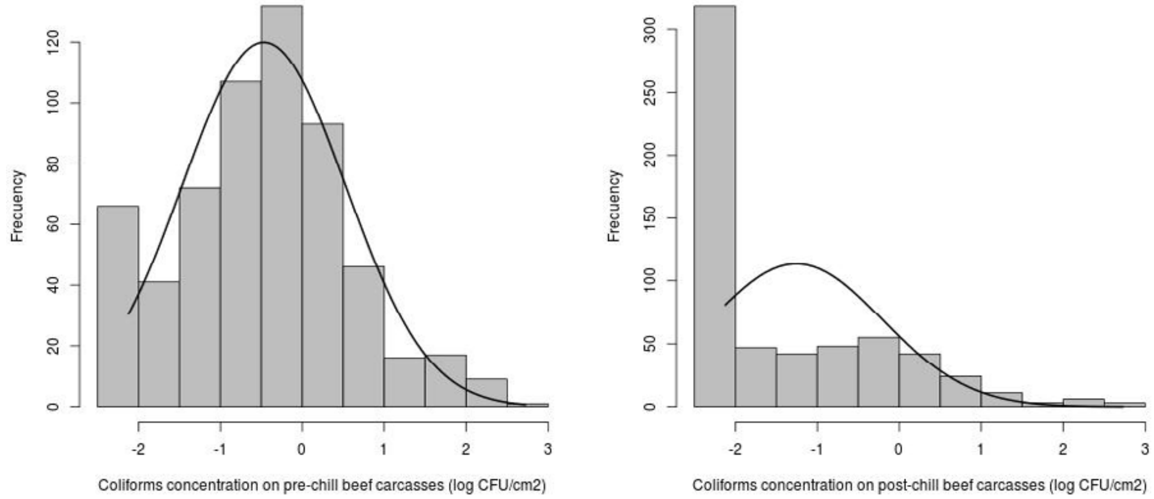

FIGURE 1. Histograms of frequency of the concentration of total coliforms on beef carcasses before (left) and after chilling (right) sampled from eight Irish abattoirs. Zero counts were replaced by the microbiological limit of enumeration of -2.0 log CFU/cm^2. Lack of fit to a normal distribution ($p<0.001$) is graphically shown

Due to the high level of zero counts (equivalent to less than −2.0 log CFU/cm^2), the application of the Box–Cox method (Peltier et al., 1998) failed to approximate the data to a normal distribution. It may be argued that inferential statistics in these cases can be conducted using non-parametric statistics (e.g. Mann–Whitney, Kruskal–Wallis tests) or even outcome categorisation. However, although they may be acceptable in certain applications, these techniques normally result in loss of information, which is detrimental in prediction and modelling applications. Thus, it is clear that, while logarithmic transformation to induce data normality can be suitable for bacterial counts of high occurrence such as mesophile or total viable counts (Gonzales-Barron and Butler, 2011a), this approach may be inappropriate for bacterial counts of lower occurrence such as coliforms or pathogens. Gonzales-Barron and Butler (2011a,b) and Gonzales-Barron et al. (2012; 2013) have demonstrated that the Poisson-gamma distribution is much more suitable than the lognormal distribution to represent low microbial counts and even highly clustered microbial data consisting of a large proportion of zero counts. Thus, to carry out inferential and predictive statistics without incurring loss of information, an alternative approach is to conduct count data Poisson-gamma regression models. For the fitting of these count data models, the coliforms counts cannot be expressed in terms of a continuous variable (log CFU/cm^2). A discrete response variable must be used, which is defined as the number of colonies counted on a Petri dish originated from the 1-ml aliquot (Y in CFU).

The simplest Poisson-gamma regression model (Model 1), with a significant dispersion parameter k (Table 1), suggests that the over-dispersion of the plate count data is indeed accounted for by the Poisson-gamma distribution. It also evidences the significant decreasing effect that chilling has on the coliforms viability by an overall factor of 0.133 ($\exp(\beta_1)=\exp(-2.017)$). Nevertheless, there is some uncertainty around this value due to the significant heterogeneity among the batches nested in the abattoirs, represented by a standard deviation σ_u of 1.99 (Table 1). The fluctuation found batch to batch in the reducing effect of chilling is not unexpected as there are many factors contributing to the variability in the detected coliforms: hygiene of the abattoirs, size of the abattoirs, type and capacity of the chilling systems, logistics, carcass surface dryness, proximity of contaminated carcasses, etc.

The objective of the second Poisson-gamma model was not to estimate the effect of chilling but to test whether there is any correlation between the intercept β_0 and the slope β_1 of the logarithmic function of the mean concentration λ (Equation 7). For this, random effects for the batches (nested within abattoirs) were placed in both parameters. The significant coefficient of correlation ρ=-0.7 suggested that as the initial (pre-chill) contamination increases ($\exp(\beta_0)$), the reducing effect of chilling ($\exp(\beta_1)$) on the viability of coliforms tends to be weaker. The association can be clearly visualised in Figure 2, where each of the markers represents the exponential values of the paired random effects ($\exp(v_{b(a)})$, $\exp(u_{b(a)})$) of a batch. Notice that in Model 2, both β_0 and β_1 were not significantly different from zero, and hence they were dropped.

When a log-linear predictor for the dispersion parameter k was introduced (Model 3), the goodness-of-fit measure BIC was considerably improved (in relation to Model 1) since chilling significantly affected k. Thus, the dispersion parameter for the pre-chill carcasses group was lower (4.6) and significantly different from the one for the post-chill group (9.7; Table 1). As explained in Gonzales-Barron and Butler (2011a), the higher the proportion of zeros in a data set, the more skewed the Poisson-gamma distribution becomes, and hence, the higher the dispersion parameter.

TABLE 1. Mean parameter estimates and standard errors (in brackets) of the four random-effects Poisson-gamma regression models fitted to the plate count data of total coliforms on Irish beef carcasses

	Model 1	**Model 2**	**Model 3**	**Model 4**
No. parameters	**3**	**4**	**5**	**6**
Parameters				
β_0	1.820***	0	1.816***	1.816***
	(0.103)		(0.089)	(0.017)
β_1	-2.017***	0	-1.950***	-1.834***
	(0.470)		(0.465)	(0.061)
k	6.300***	5.121***	-	-
	(0.243)	(0.204)		
α_0	-	-	1.533***	1.533***
			(0.048)	(0.027)
α_1	-	-	0.744***	0.810***
			(0.081)	(0.037)
Random effects				
σ^2_u	3.961*	3.626*	3.786*	2.922***
	(1.328)	(1.195)	(1.318)	(0.058)
σ^2_v	-	7.904*	-	0.599***
		(2.602)		(0.009)
ρ	-	-0.698*	-	0
		(0.120)		
Goodness-of-fit				
BIC	11245	11034	11159	11076
Other estimates				
λ **before**	6.171	-	6.148	6.148
(CFU/cm^2)	(0.636)		(0.544)	(0.106)
λ **after**	0.822	-	0.875	0.982
(CFU/cm^2)	(0.376)		(0.400)	(0.061)
k **before**	-	-	4.633	4.633
			(0.220)	(0.123)
k **after**	-	-	9.748	10.415
			(0.639)	(0.241)

We could then generalise that, should the effect of chilling be different batch to batch – as already suggested by the previous models, the proportion of zero counts in the pre-chill carcasses relative to the post-chill carcasses will be different batch to batch, and this will lead to different pre-chill to post-chill dispersion parameter rates batch to batch. With this in mind, in Model 4, random effects $v_{b(a)}$ were added to the slope (effect of chilling α_1) of the dispersion parameter predictive equation (Equation 9). In this Model 4, all parameters were significant and the goodness-of-fit measure BIC was considerably improved (in comparison to Model 3). No correlation was found between the random effects for the chilling effect on the mean microbial concentration ($u_{b(a)}$) and the random effects for the chilling effect on the dispersion parameter ($v_{b(a)}$). Overall, the parameter estimates of Model 4 did not differ considerably from those of Model 3, although the additional random effects of Model 4, accounting for the batch-to-batch heterogeneity in the effect of chilling on the dispersion parameter, decreased the standard errors of the parameter estimates (Table 1). It can be said that the chilling operation reduces the viability of total coliforms on beef carcasses by a mean factor of 0.16 (exp(-1.834)). On average, before chilling the beef carcasses sampled from the Irish abattoirs presented a concentration of coliforms of 6.15 CFU/cm^2, and after chilling, this concentration dropped significantly to 0.98 CFU/cm^2.

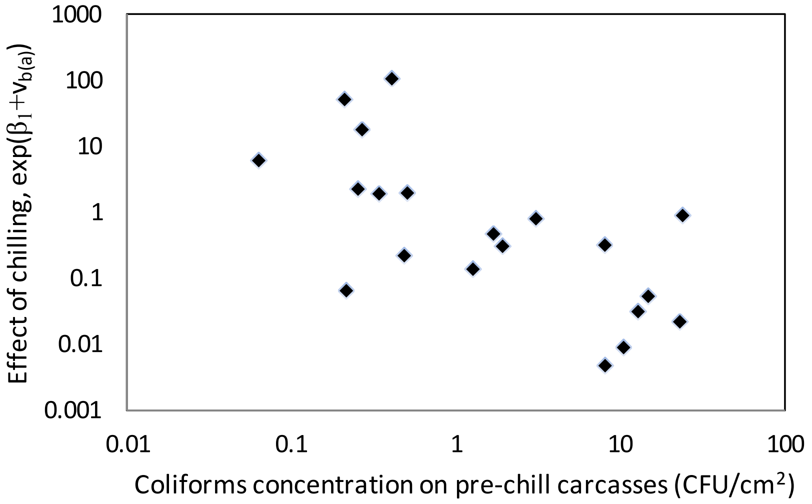

FIGURE 2. Correlation between the initial concentration of coliforms on a batch of pre-chill beef carcasses and the effect of chilling (ρ=-0.7)

The Poisson-gamma regression framework constitutes a neat approach to tackle statistical analysis of low counts microbial data because it does not make use either of the ordinary log-transformed data or the imputation of the lowest limit of enumeration for the zero counts, and instead it uses directly the observed number of colony forming units and takes into analysis all the zero counts.

4. CONCLUSIONS

This work introduced an alternative framework based on the Poisson-gamma regression to conduct inferential statistics on plate count data from microorganisms of low recovery. As an illustration of this type of count data models, we estimated the effect that chilling has on the mean counts of the total coliforms on beef carcasses as well as its between-batch variability. Such a neat approach has a potential utility in risk assessment modelling of pathogens and in the design of sampling plans.

5. REFERENCES

1. Gill, C., Deslandes, B., Rahn, K., Houde, A., Bryant, J., (1998). ‘Evaluation of the hygienic performances of the processes for beef carcass dressing at 10 packing plants’. Journal of Applied Microbiology 84, 1050–1058.
2. Gonzales-Barron, U., Zwietering, M., and Butler, F., (2013). ‘A novel derivation of a within-batch testing regime based on a Poisson-gamma model characterising low microbial counts in foods’. International Journal of Food Microbiology, 161(2): 84-96.
3. Gonzales-Barron, U., Lenahan, M., Sheridan, J.J., and Butler, F., (2012). ‘Use of a Poisson-gamma model to assess the performance of the EC process hygiene criterion for Enterobacteriaceae on Irish sheep carcasses’. Food Control, 25: 172-183.
4. Gonzales-Barron, U., and Butler, F., (2011b). ‘Characterisation of within-batch and between-batch variability in microbial counts in foods using Poisson-gamma and Poisson-lognormal regression models’. Food Control, 22: 1268-1278.
5. Gonzales-Barron, U., and Butler, F., (2011a). ‘A comparison between the discrete Poisson-Gamma and Poisson-Lognormal distributions to characterise microbial counts in foods’. Food Control, 22: 1279-1286.
6. Gonzales Barron, U., Kerr, M., Sheridan, J., and Butler, F., (2010). ‘Count data distributions and their zero-modified equivalents as a framework for modeling microbial data with a relatively high occurrence of zero counts’. International Journal of Food Microbiology, 136: 268–277.
7. Peltier, M.R.,Wilcox, C.J., Sharp,D.C., (1998).Application of the Box–Cox data transformation to animal science experiments. Journal of Animal Science, 76: 847–849.