Food Control 37 (2014) 385–394

Contents lists available at ScienceDirect

Food Control

journal homepage: www.elsevier.com/locate/foodcont

Conducting inferential statistics for low microbial counts in foods using the Poisson-gamma regression

Ursula Gonzales-Barron [a, *], Vasco Cadavez [a], Francis Butler [b]

[a] CIMO Mountain Research Centre, ESA School of Agriculture, Polytechnic Institute of Braganza, Campus de Santa Apolónia, Apartado 1172, Braganza 5301-854, Portugal
[b] UCD School of Biosystems Engineering, University College Dublin, UCD Belfield, Dublin 4, Ireland

ARTICLE INFO

Article history:
Received 22 May 2013
Received in revised form
24 August 2013
Accepted 14 September 2013

Keywords:
Carcass
Chilling
Coliforms
Negative binomial
Zero-inflated
Plate counts

ABSTRACT

Mixed Poisson distributions have been shown to be able to represent low microbial counts more efficiently than the lognormal distribution because of its greater flexibility to model microbial clustering even when data consist of a large proportion of zero counts. The objective of this study was to develop an alternative modelling framework for low microbial counts based on heterogeneous Poisson regressions. As an illustration, Poisson-gamma regression models were used to assess the effect of chilling on the concentration of total coliforms from beef carcasses ($n = 600$) sampled at eight large Irish abattoirs. Three Poisson-gamma and three zero-modified (hurdle and zero-inflated) models were appraised with a series of random-effects variants in order to extract any variability in microbial mean concentration, dispersion and/or proportion of zero counts. Models were compared and validated in their ability to predict the coliforms counts on carcasses after chilling. In all five test batches, the hurdle Poisson-gamma distributions predicted the observed post-chill counts closer than the Poisson-gamma distributions. This is justified by the better capacity of the hurdle model to represent a higher proportion of zero counts, which were in fact observed in the post-chill batches. Thus, with a coded variable (pre-chill/post-chill) as treatment, and extracting the significant variability of batches nested in abattoirs for the coliforms mean concentration ($\sigma^2_u = 2.68$), the dispersion measure ($\sigma^2_v = 2.39$) and the probability of zero counts ($\sigma^2_w = 0.89$), the validated hurdle Poisson-gamma model confirmed that chilling has a decreasing effect on the viability of coliforms from beef carcasses, and that the concentration is reduced by an average (pre-chill to post-chill) factor of 2.2 (95% CI: 2.15–2.24) at batch level. The model also indicated that chilling increases the odds of producing a zero count from a carcass swab in about 13.5 times, and that the higher the coliforms concentration in a batch, the weaker the effect that chilling has to reduce such contamination on the beef carcasses.

1. Introduction

In the evaluation of microbiological quality of foodstuffs, bacterial concentration is conventionally expressed in terms of log CFU cm^{-2} or g^{-1}. Logarithmic transformation is believed to approximate data normality, which is fundamental for the application of inferential statistical data analysis based on the Gaussian distribution. This assumption leads 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 inserted. 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 overestimated. To avoid this, Busschaert, Geeraerd, Uyttendaele, and Van Impe (2011) developed a censored data method to fit a lognormal distribution to plate counting data with zero counts, which later on Pouillot, Hoelzer, Chen, and Dennis (2013) adapted for fitting lognormal and zero-inflated lognormal distributions to Most Probable Number (MPN) data. On the other hand, Gonzales-Barron & 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 in the characterisation of low microbial counts and even for highly clustered microbial data consisting of a large proportion of zero counts (non-detections). Whereas the lognormal distribution cannot naturally take in zero values, the Poisson-lognormal distribution greatly overestimates the mean microbial concentration when it attempts to fit in a high

* Corresponding author. Tel.: +351 273 303325.
E-mail address: ubarron@ipb.pt (U. Gonzales-Barron).

0956-7135/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.foodcont.2013.09.032

frequency of zero counts (Gonzales-Barron & Butler, 2011a). As a consequence, parametric statistical analyses, such as ANOVA or regression models, should only be performed on bacterial counts of high occurrence (such as mesophile or total viable counts), for which a logarithmic transformation can mostly bring about an approximate normal distribution. For bacterial counts of lower occurrence (such as some hygiene indicators and pathogens) that can be better described by a Poisson-gamma distribution, the alternative modelling framework of the Poisson-gamma regression should be explored.

The heterogeneous Poisson models and their zero-inflated variants correspond to a family of flexible count data distributions that can handle over-dispersion (i.e., variance of the observed count data exceeds the mean), which is a condition caused by microbial heterogeneity or clustering. A heterogeneous Poisson model loosens the Poisson restriction (whereby the nominal variance equals the mean) 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. However, with some types of data, over-dispersion may also stem from a high percentage of zero counts, for which the variance function of the Poisson-gamma model may be insufficient. In this case, the zero-modified Poisson-gamma regression models, such as zero-inflated and hurdle, may be more convenient (Gonzales-Barron, Kerr, Sheridan, & Butler, 2010). Therefore, the objective of this work was to introduce a count data regression framework to conduct inferential statistics on microbial counts that do not approximate to a normal distribution after logarithmic transformation due to clustering and the relatively high proportion of zero counts. As an illustration, we will use the proposed Poisson-gamma and zero-modified Poisson-gamma regression models to assess whether there is an effect of chilling on the total coliforms recovered from beef carcasses; and, if so, to quantify the expected chilling effect at batch level. For this aim, a number of random-effects models were evaluated and compared in terms of goodness-of-fit and predictability.

2. Methodology

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

Eight beef export abattoirs, with a throughput of at least 30,000 cattle per annum each, located in the south, east and west of Ireland, were visited during 2007–2008 to obtain a representative sample of cattle being slaughtered throughout the country. Four of the abattoirs were each visited three times and the remaining four on two occasions. During each visit, 30 animals were randomly sampled throughout the day to obtain a representative sample of the cattle being presented for slaughter. After washing, at the end of the slaughter line, the two carcass sides of an animal to be sampled one before and one after chilling, were identified. Polyurethane sponges (Sydney Heath, Stoke on Trent, UK) were cut for use as carcass swabs (150 cm^2), which were pre-soaked in 5 ml of Universal Quenching Agent (UQA). This solution contained 1.0 g peptone, 1.0 g sodium thiosulphate, 5.0 g Tween 80 and 0.7 g lecithin, in a litre of distilled water, at pH 7.0. The UQA solution was used because the levels of chlorine in the carcass wash water varied depending on the abattoir being visited. Swabs were placed in stomacher bags (Sarstedt, Numbrecht, Germany) and autoclaved at 115 °C for 10 min. Carcasses were swabbed using the method described by Lasta, Rodriguez, Zanelli, and Margaria (1992), which involved uniformly swabbing the entire outer surface of a carcass from the hindquarter to the forequarter. The extent of carcass swabbed was estimated according to Lasta et al. (1992) and annotated for each animal. Areas swabbed ranged between 8900 and 11,200 cm^2 with a mean of ~10,000 cm^2. A single sponge was used to swab each side of the carcass.

Two hundred millilitre of Buffered Peptone Water (BPW) (Oxoid, Basinstoke, UK) was added to each sponge and homogenised in a Colworth stomacher (Model 400, Seward, London, UK) at 200 rpm for 1.0 min. A 15-ml aliquot of the homogenate was poured into 30 ml sterile polystyrene screw cap containers (Sterilin, Stone, UK). To determine total coliforms, 1-ml volumes of the neat homogenate were dispensed into each of two Petri dishes and over-poured with Chromocult coliform agar (Merck, Darmstadt, Germany) and incubated at 37 °C for 24 h. Total coliforms were identified as salmon-coloured colonies. Although the concept of microbiological limit of enumeration is not used in the Poisson-gamma modelling framework, it is worthy to mention that, for the present protocol, this value was approximately −2 log CFU/cm^2, assuming a mean carcass swab area of 10,000 cm^2 (If one colony was counted in one of the two plates, the threshold would correspond to 1 CFU/(2×50 cm^2)).

2.2. Poisson-gamma 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 carcass surface to swabs and from the swabs to the homogenate (i.e., perfect recovery); (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,

$$\Pr(Y_i) = \frac{\exp(-\mu_i) \times \mu_i}{Y_i!} \quad (1)$$

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

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 of the carcass sample 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 consisting of an intercept β_0 and the chilling effect β_1. In our case, we wish to quantify the effect of the chilling treatment β_1; thus, the only covariate was a coded variable X (0 as pre-chill, 1 as post-chill),

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

Within a batch, carcasses do not share the same true unknown microbial concentration λ_i (CFU/cm^2) due to the heterogeneity in contamination arising from systematic errors along the slaughter line and/or random sources of variability. Thus, the basic Poisson regression model was generalised by including a dispersion parameter to accommodate the heterogeneity in the count data. A generalised Poisson distribution lets the expected microbial concentration λ_i be a function also of some random variable e_i (Hinde & Demetrio, 1998),

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