Transnational University Limburg is a unique collaboration of two universities in two countries: the University of Hasselt and Maastricht University.

Universiteit Hasselt | Campus Hasselt | Martelarenlaan 42 | BE-3500 Hasselt
Universiteit Hasselt | Campus Diepenbeek | Agoralaan Gebouw D | BE-3590 Diepenbeek

2012
2013

FACULTY OF SCIENCES

Master of Statistics: Epidemiology & Public Health Methodology

Masterproef

Understanding the effect of raw materials on some characteristics of meat emulsions for pet food production

Promotor :
dr. Francesca SOLMI

Cabrale Nango

Master Thesis nominated to obtain the degree of Master of Statistics , specialization Epidemiology & Public Health Methodology

2012
2013

FACULTY OF SCIENCES

Master of Statistics: Epidemiology & Public Health Methodology

Masterproef

Understanding the effect of raw materials on some characteristics of meat emulsions for pet food production

Promotor :
dr. Francesca SOLMI

Cabrale Nango

Master Thesis nominated to obtain the degree of Master of Statistics , specialization Epidemiology & Public Health Methodology

Interuniversity Institute for Biostatistics and Statistical Bioinformatics

Universiteit Hasselt

UNDERSTANDING THE EFFECT OF RAW MATERIALS ON SOME CHARACTERISTICS OF MEAT EMULSIONS FOR FOOD PRODUCTION.

By: *Nango Cabrale*

Supervisor: *Dr. Francesca SOLMI*

Thesis submitted in partial fulfilment of the requirements for the degree of Master of Science in Statistics: Epidemiology and Public Health Methodology

February 2013

CERTIFICATION

This is to certify that this report was written by Nango Cabrale under our Supervision.

Nango Cabrale

Student

……………………..

Signature

……………………….

Date

Dr. Francesca SOLMI

Supervisor

………………………

Signature

……………………………..

Date

ACKNOWLEDGEMENT

First and foremost I am grateful to the Almighty God for the health and intercessions bestowed upon me to complete this master programme.

I would like to thank all my lecturers whose knowledge helped me immensely in the realization of this work.

I wish to express my sincerest gratitude to my thesis advisor, Dr. Francesca SOLMI, for her insightful guidance, constant encourage, diligence, expertise and pertinent questions during discussions in this area which lead to the successful completion of the thesis.

My gratitude goes to my friend Nico Ntellah alongside all the CamUHasselt fellows and Dschang University course mates (Sylvestre, Paoussi, Amerine, Lydienne) for their advises and moral support. My special thanks to D'Espagne Sonna and Anne Momo for their cares and moral support.

My very special thanks go to my late father (Tatsaou Pierre) and my mother (Fokeng Veronique), brothers (Bertin Assontsa, Alain, Christian, Maurice, Mr Folefack Jerome), sisters (Adéline and Marie) for their constant support, encouragement, love and without whom it would not be feasible to complete this master of statistics programme.

Abstract

This study investigates the effect of raw materials on some characteristics of meat emulsion for food production; seek to understand how it is possible to improve Hardness and Resilience of the final product after cooking by changing the percentages of the raw materials. The data set is a subset of 129 observations corresponding to different combinations of percentages of 8 raw materials. To accomplish our task in this research a mixture regression model was developed due to their advantages over the classical models such as Tobit model in model fit when data are generated from a two step process. Additionally, the model is shown to allow for flexibility in distributional assumption. However, zero-inflated gamma and zero-inflated log normal models were used to evaluate the effect of raw material on Hardness and Resilience of the end product and to account for both the presence of zeroes values and the positive skewness in these outcomes variable. The results showed that the zero-inflated gamma model clearly models and predicts with more accuracy than the zero-inflated log normal model for Resilience outcome whereas the zero-inflated log normal predicted better than the zero inflated gamma for Hardness outcome based on the AIC.

Contents

List of Tables

List of Figures

1. Introduction

Data analysts are often confronted with the issue of choosing the right model because a variety of models can be derived for a single dataset. Thus the model building process based on sound scientific principles is important (Burnham & Anderson, 2002). In specifying a statistical model in the GLM framework, there is a need to carefully choose the random and systematic components, as well as an appropriate link function (Keele, 2008).

This project looks into meat emulsions for food production and the effect of raw materials on some of its characteristics. Meat emulsions are finely comminuted meat mixtures composed of water, protein, fat, salt and small amounts of other ingredients. This meat emulsions which are comminuted meat products, are well known in the food industry and are widely used in the production of products such as balogna, frankfurters and other sausage products. Such meat emulsion products are prepared by mixing, chopping, and emulsifying a mixture of raw meat materials, such as lean skeletal beef and pork and meat by-products, with ice, salt, spices and curing salts in such a manner as to produce an emulsion which contains fine fat particles coated with protein dissolved from the meat ingredients. The resulting meat emulsion is then stuffed into suitable casings, which serve as processing molds, and are heated at increasing temperatures of from 55°C. to 77°C. for extended periods of time which may vary between about 1 to 8 hours or more, depending on the volume of the meat emulsion being processed. Upon such heating, the protein in the meat emulsion coagulates or sets solid and entraps the fat particles in the protein matrix thereby forming a firm meat emulsion product. Such meat emulsion products are a uniform homogeneous mass which contain no discrete particles of meat and retain the shape of the casing when set.

In recent years in order to reduce the cost of certain food products to consumers, there has been an increasing demand for chunky food products which resemble chunks or pieces of natural meat in appearance, texture and physical structure, and which may be used as a partial or complete replacement for the more expensive natural meat chunks in food products such as stews, pot pies, casseroles, canned foods and pet food products. Chunky meat products are highly desirable in both human foods and pet foods both from aesthetic quality and consumer appeal. Because of this desirability and the high ingredient cost of natural meat chunks, there is a need for replacement of such expensive natural meat chunks in foods with more economical chunky products which simulate natural meat chunks in shape, appearance and texture, and which retain their shape, appearance and texture when subjected to commercial canning and retorting procedures.

Heretofore, efforts directed to providing such simulated natural meat chunks have been directed to producing such products from vegetable protein sources using extrusion-expansion techniques. Although the products of such extrusion-expansion procedures have met with some acceptance in the food industry, their use has been limited primarily to use as meat extenders. Since such products lack the appearance and texture of natural meat they are not generally suitable for use as full substitutes for meat. Similarly, meat emulsion products produced by conventional procedures, which are in the form of a uniform, homogeneous mass, lack the structure, texture and appearance of natural meat chunks and are not suitable for use in applications in which the use of simulated natural meat chunks is desired.

The present invention provides for the production of a meat emulsion product in the form of distinct chunks or pieces having a plurality of juxtaposed, manually separable meat-like layers resembling a chunk of natural meat in appearance, texture, and consistency. The meat emulsion chunks of this invention are suitable for use as a partial or complete replacement for more expensive natural meat chunks in both human foods and animal foods, and retain their integrity and shape when subjected to commercial canning and sterilization procedures such as those required in the production of canned high moisture food products.

Much progress has been made over the last fifty years in understanding what the texture of a particular food is, and how it can be measured, specified and controlled. Knowledge of the textural properties of processing meat emulsion is crucial to ensuring product acceptability; measurement, control, and optimization of these properties through judicious selection of varieties and control of unit operations results in products that the consumer prefers. In food consumption as mentioned above, appearance, flavor and texture of the obtained end product are the three major acceptability factors because they can impart enjoyment of the food. If these attributes do not meet consumer expectations, the food will not be consumed and the customer is unlikely to purchase that particular brand of product again. Additionally, consumer products succeed in the marketplace in part because their "textural characteristics" are pleasing to customers. This is certainly true with food products but it also applies to cosmetics, pharmaceuticals, packaging, industrial materials and even adhesive type materials.

Appearance (color, size, shape) is based on the optical sense and is a response of the eye to the light reflected from or transmitted through the food. Flavor is the body's response to a chemical impact and is sensed in two locations: 1) the olfactory organ in the nose (aroma or smell), and 2) the taste buds in the tongue (taste), these are called the chemical sense. Texture considered here as Hardness (soft, compact, hard, defined as force necessary to attain a given deformation) and Resilience (tender, chewy, tough, defined as measurement of how a sample recovers from deformation in relation to speed and forces derived) of the end product is sensed primarily in the mouth, on the lips, teeth, gums and tongue, although some texture notes can be sensed by other part of the body, such as the hand. Textural perception occurs directly through the tactile (touch) and kinesthetic (movement) senses, and indirectly through the senses of vision and hearing. In contrast to color and flavor, there are no specific sensory receptors for texture. Texture is an important quality attribute in almost all foods (e.g canning foods), and is most important in foods that are bland in flavor or have the characteristics of freshness or crunchiness.

Interest in what comprises texture and how it is measured and controlled is driven by two major concerns: 1) imparting pleasure just before and during mastication and 2) economics. People are prepared to pay a higher price for food or any other manufactured product when the texture is "just right".

A good example of this human propensity for textures that please can be found in the case of meat from American supermarkets, where different cuts of beef typically range in price from less than three dollars to more than sixteen dollars per kilogram. This wide range in price is largely the result of texture quality: consumers are prepared to pay a high price for tender meat and expect to pay low price for tough or dry meat. Considering the many millions of kilograms of beef or meat emulsion consumed each year, it becomes obvious that economic factors are a great driving force to achieving desirable textures in beef and other foods such as meat emulsion. Almost all researchers agree that "texture" is a sensory attribute and that a number of textural properties exist. The International Organization for Standardization defines texture as "all the mechanical, geometrical and surface attributes of a food product perceptible by means of mechanical, tactile, and, where appropriate, visual and auditory receptors". Those physical properties of foods that are not sensed by the body (and there are many) should not be described as texture. There are often good reasons for measuring non-textured physical properties, but they should not be confused with textural properties.

The rationale of this project therefore was to understand the effect of raw materials on Hardness and Resilience of meat emulsion for food production; understand how it is possible to improve these characteristics (here Hardness and Resilience) of the final product after cooking by changing the percentages of the raw materials.

To accomplish this task, zero-inflated log normal (ZILN) and zero-inflated gamma (ZIG) models derived from a generalized mixture models were used.

The data and methodology are presented in section 2, results; discussions and conclusion in sections 4 and 5 respectively while section 3 presents the model selection. Section 6 has some limitations and recommendations.

2. Data and Methods

2.1. Data Description

The dataset at disposal is a subset of 129 observations corresponding to different combinations of percentages of eight raw materials: Raw material 1, 2, 3, 4, 5, 6, 7 and 8. The two outcomes of interest in the analysis were Hardness and Resilience of the obtained final product after cooking. All the predictors and outcomes variables are semicontinuous. The 129 combinations of raw materials summed to 100% inducing multicollinearity problem in the data. The Variance Inflation Factor (VIF) was computed and was found to be very high (VIF = 6029401 for variable 'raw material 8' and greater or equal to 5873710 for the other covariates). In order to solve this problem, variable "raw material 8" was dropped from the analysis. The recorded covariates and outcomes are presented in Table 1.

2.2. Exploratory data Analysis (EDA)

In order to gain insight into the data set, exploratory data analysis was conducted. Given the nature of the data, tables, density functions, descriptive statistics, Pearson correlation and scatter plots matrix were used to display the data. Investigation of these graphical representations conveys possible relationships of interest within the dataset which were investigated in the model selection section in more details.

2.3. Statistical Analysis

Dependent variables with many zeroes values have long been a complexity associated with micro data sets. In some applications, the response variable can take any nonnegative value but has positive probability of a zero outcome. We refer to a variable as semicontinuous when it has a continuous distribution except for a probability mass at 0 such as Hardness of 0g and Resilience of 0joule/m3. Semicontinuous data are common in many areas. For example, when each observation is a record of the total rainfall in the previous day, many days have no rainfall. In a study of annual medical costs, a portion of the population has zero medical expense. With semicontinuous data, unlike censored data, the zeros represent actual response outcomes. The most common occurrences are found in consumption and production data. Regarding consumption, households typically do not purchase all of the goods being evaluated in every time period and some of these households spend nothing on a certain commodity during the period of investigation. Similarly, a study evaluating the percentage of mortality rates in a cattle production yields will likely have outcomes with high percentage of mortality or with zero percentage of mortality rates. In both cases, ordinary least squares parameter estimates will be biased when applied to these types of regressions (Amemiya, 1984).

The seminal work by Tobin (1958) was the first to recognize this bias and offer a solution that is still quite popular today. The univariate Tobit model is extended, under a mild set of assumptions, to include multivariate settings (Amemiya, 1974; Lee, 1993). While empirical applications in univariate settings are discussed by Amemiya (1984), multivariate applications are becoming more frequent (Belasco, Goodwin and Ghosh, 2007; Chavas and Kim, 2004; Cornick, Cox and Gould, 1994; Eiswerth and Shonkwiler, 2006). While the Tobit model has had a large impact on modeling censored dependent variables, it is not without limitations. The two major assumptions made by the Tobit model in its original derivations included the assumption of normality and the point that both the observable and unobservable variable levels come from the same distribution. The assumption of normality has made the Tobit model inflexible to data generating processes outside of that major distribution (Bera et al., 1984). Additionally, Arabmazar and Schmidt (1982) demonstrate that random variables modeled by the Tobit model contain substantial bias when the true distribution is non-normal and has a high degree of censoring. Inconsistent estimation results arise when residuals are positively skewed. Additionally, maximum likelihood estimation becomes complicated with a system of equations when the zeroes values occurs in multiple equations because of the problem of integrating more than three integrals in the likelihood.

Another important model that is also used to characterize observations containing many zeros is the zero-inflated models. These models have been rarely used in statistics. The Poisson distribution is commonly placed into a zero-inflated framework and is appropriately called the zero-inflated Poisson (ZIP) model. The advantage to using this type of model is again that it recognizes that decisions or production output processes are part of a two step process. In the other hands, Hardness and Resilience of meat emulsion, which are provided in this data set, provide valuable insights into the profitability and performance of food production companies. Additionally, the measures of these outcomes may be more accurately characterized by a mixture model that takes into account their positive skewness as well as allowing zeroes and non-zeroes observations to be modeled independently. A zero-inflated specification (zero inflated log normal(ZILN) and zero inflated gamma (ZIG) models) is used rather than other mixture specifications, such as Hurdle model to more accurately capture measure of Hardness and Resilience of the final product after cooking. One drawback from using the Tobit Model is that the optimization routine necessary to estimate all parameters may take quite long to converge. The next essay will then focus on a Zero inflated (ZILN and ZIG) regression model that may work to shrink the computational burden from estimating so many parameters, as well as improve estimation efficiency.

Having noted the nature of the responses variables (Hardness and Resilience in our case) as always non-negative containing zeroes values and positively skewed various, models for continuous and positive skewed dataset such as lognormal and gamma were considered. To compare model fits, we use the classical computation of Akaike's Information Criteria (AIC) (Akaike, 1974) as discussed in the paper by Eric J. Belasco and Sujit K. Ghosh (2008) and negative log-likelihoods (actually -2* the log-likelihood) values as discussed in the paper by Mathew Flynn et al (2009). Likelihood ratio test was not applied since it can only be used to compare a nested model (Verbeke and Molenberghs, 2008/2009) which is not the case here. We consider the use of a mixture model to characterize dependent variables with zeroes values as an alternative to the Tobit model. Its major advantages include the flexibility in distributional assumptions and an increased efficiency in situations involving a high degree of zeroes. We derive the zero-inflated log-normal and gamma model from a generalized mixture model.

2.3.1. Generalized Mixture Models

In general, mixture models characterize dependent variables with zeroes values (here Hardness and Resilience) as a function of two distributions (Y = VB). First, B measures the likelihood of zero or positive outcomes, which have been characterized in the literature using Bernoulli and Probit model specifications. Then, the positive outcomes are independently modeled as V. A major difference between the mixture and Tobit model is that unobservable, censored observations are not directly estimated. A generalized mixture model can be characterized as follows:

$$f(y|\theta) = 1 - \rho(\theta) \qquad y = 0$$

$$= \rho(\theta) g(y|\theta) \quad y > 0 \tag{1}$$

Where $\int_0^\infty g(y|\theta)dy = 1 \quad \forall\theta$. This formulation includes the standard univariate Tobit model when $\theta = (\mu, \sigma), \rho(\theta) = \Phi\left(\frac{\mu}{\sigma}\right)$, and $g(y|\theta) = \frac{\phi\left(\frac{y-\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)} I(y > 0)$. Notice that in Log-normal and Gamma zero-inflated specifications to follow, ρ is modeled independently of mean and variance parameter estimates, making them more flexible than the Tobit model. Next, we develop two Univariate zero-inflated models that include covariate variables such as raw materials, which then can be extended to allow for multivariate cases. Since only the positive outcomes are modeled through the second component, the log of the dependent variables can be taken. Taking the log of these variables works to symmetrize the dependent variables that were originally positively skewed. Using a log-normal distribution for the V random variable and allowing ρ to vary based on the conditioning variables, we can transform the basic zero-inflated model into the following form that can be generalized to include continuous distributions. We start by deriving the normal distribution to model the logarithm of the dependent variable outcomes, also known as the log-normal distribution, of the following form:

$$f(y_i|\beta,\alpha,\delta) = 1-\rho_i(\delta) \qquad for\ y_i = 0$$

$$= \rho_i(\delta)\frac{1}{y_i}\phi\left(\frac{\log(y_i)-X_i'\beta}{\sigma_i}\right) \qquad for\ y_i > 0 \qquad (2)$$

Where

$$\rho_i(\delta) = \frac{1}{1+\exp(X_i'\delta)} \qquad (3)$$

$$\sigma_i^2 = \exp(X_i'\alpha) \qquad (4)$$

Which guarantees σ_i^2 to be positive and $\rho_i(\delta)$ to be between 0 and 1 for all observations and all the parameters values. $\phi(.)$ denotes the probability density function of a standard normal distribution with mean zero and variance unity; K is equal to the number of conditioning predictor variables or covariates; $\delta, \beta\ and\ \alpha$ are (Kx1) vectors of regression coefficients; X_i' is a (1×K) vector of predictor variables (raw materials in our case); $\rho_i(\delta)$ is the logit link function; σ_i^2 is the conditional variance and y_i is the responses variable (Hardness and Resilience). Notice that this specification is nested within the generalized version in equation (1) where $g(y|\theta)$ is a log-normal distribution and $\theta = (\delta, \beta, \alpha)$. This model will be denoted by ZILN (ρ, μ, σ^2) where ρ denotes the link function (e.g. logit, probit, etc.), μ denotes the mean function on the logarithmic scale and σ^2 denotes the variance function on the logarithmic scale of the positive part of Hardness and/or Resilience.

In addition to deriving a zero-inflated log-normal distribution, we will also derive a zero-inflated Gamma distribution to demonstrate the flexibility of the zero-inflated regression models and perhaps improve upon modeling a variable that possesses positive skewedness. Within a univariate framework, the sampling distribution can be easily changed by deriving V as an alternative distribution in much the same way as equation (2). Following is the specification for the zero-inflated Gamma distribution, where V is distributed as a Gamma distribution where the shape parameter is λ_i, and η_i is the rate parameter.

$$
\begin{aligned}
f(y_i|\lambda_i, \eta_i, \delta) &= 1 - \rho_i(\delta) && for \ \ y_i = 0 \\
&= \rho_i(\delta) \, \frac{y_i^{\lambda_i - 1} e^{-\eta_i y_i}}{\Gamma(\lambda_i)} {\eta_i}^{\lambda_i} && for \ \ y_i > 0 \qquad (5)
\end{aligned}
$$

This function can be reparameterized to include the mean of Gamma, μ by substituting $\lambda_i = \mu_i \eta_i$, where $\eta_i = e^{(x_i' \varphi)}$ and $\mu_i = e^{(x_i' \gamma)}$. Within the Gamma distribution specification, the expected value and corresponding variance can be found to be $E(y_i) = \rho_i \mu_i$ and $Var(y_i) = \rho_i(1 - \rho_i)\mu_i^2 + \rho_i \frac{\mu_i}{\eta_i}$, respectively. Both the Gamma and log-normal univariate specifications allow for a unique set of mean and variance estimates to result from each distinct set of conditioning variables.

3. Model Building

This section describes model selection procedures that were employed to choose parsimonious models. After carrying out the exploratory data analysis, we fitted a zero-inflated lognormal and gamma models due to the presence of zeroes and the positive values in the outcome variables and to account for their positive skewedness. As already mentioned in the above equations (2, 3 and 5), the parameter estimates in the zero-inflated model refer to two distinct processes. The first process includes the likelihood of a zero outcome or the one described by a log-normal or gamma distribution. This process is estimated through δ utilizing equation (3) in both log normal and gamma model. The second process includes the likelihood of the positive outcome and it is estimated through β and γ for the lognormal and gamma models respectively. Variable 'Raw material 8' was dropped from the analysis due to the multicollinearity problem and also due to the fact that individually; it was significantly negatively correlated with the outcome variables. We began by fitting the complex model with fixed effect and all the possible two and three ways interactions between covariates.

The Parameters estimates for the log normal and gamma models using log link were first used as starting values in the zero-inflated models but the algorithm did not converge and the convergence was finally achieved by using zeroes as starting values. Maximum likelihood estimation method was used to estimate parameters of both models. The optimization technique used was the Dual Quasi-Newton algorithm and the 'none' option used as the integration method for the approximation of the integral.

The backward selection methods were used to reduce the model based on the p-values for the solution of the fixed effects and interaction terms (the ones with highest p-values being eliminated first). The interactions with the largest p-value being eliminated first and the fitting procedure repeated until all the insignificant effects were eliminated from the model. Test of significance were all performed at 5% level of significance. Analysis was done using SAS 9.2 and R 2.15.0.

4. Results

This section presents EDA's findings and results of the various fitted models for the two outcomes variables (Resilience and Hardness).

4.1. Exploratory Data Analysis

As already mentioned, the 129 observations correspond to different combinations of percentages of 8 raw materials. Hardness and Resilience are the textural measured of the end product after cooking. Exploration of the covariates shows that they all have similar mean values with a slightly higher mean attributed to variable 'raw material 1'. The minimum and maximum values of all the raw materials are 0 and 1 respectively. The average Hardness and Resilience were respectively 895.6565g and 0.1408 joule/m^3 with their minimums being 0 and maximum being 2360.78 and 0.304 respectively (Table 1). Figure 1 shows the extent of correlation among the dependent variables while histograms and density plot of the dependent variables are shown in Figure 2 and Appendix Figure 1 in order to observe a clear distribution of Hardness and Resilience. Here the positively skewed nature of Hardness and Resilience of the end product are quite apparent.

Table 1: Variable description and Summary Statistics

Variables	N	Mean	Std. Dev.	Minimum	Maximum
Covariates					
Raw material 1	129	0.1265	0.2104	0	1.00
Raw material 2	129	0.1243	0.2058	0	1.00
Raw material 3	129	0.1250	0.2076	0	1.00
Raw material 4	129	0.1243	0.2067	0	1.00
Raw material 5	129	0.1243	0.2067	0	1.00
Raw material 6	129	0.1243	0.2058	0	1.00
Raw material 7	129	0.1258	0.2086	0	1.00
Responses					
Hardness	129	895.6565	477.4071	0	2360.78
Resilience	129	0.1408	0.0617	0	0.304

Table 2: Correlation between the two responses variables

Variables	Hardness	Resilience
Hardness	1.0000	0.8035 <.0001
Resilience	0.8035 <.0001	1.0000

The correlation between the responses variables and covariates was also computed and revealed a highly significant correlation between raw material 5 and the outcomes variable (Appendix Table 3).

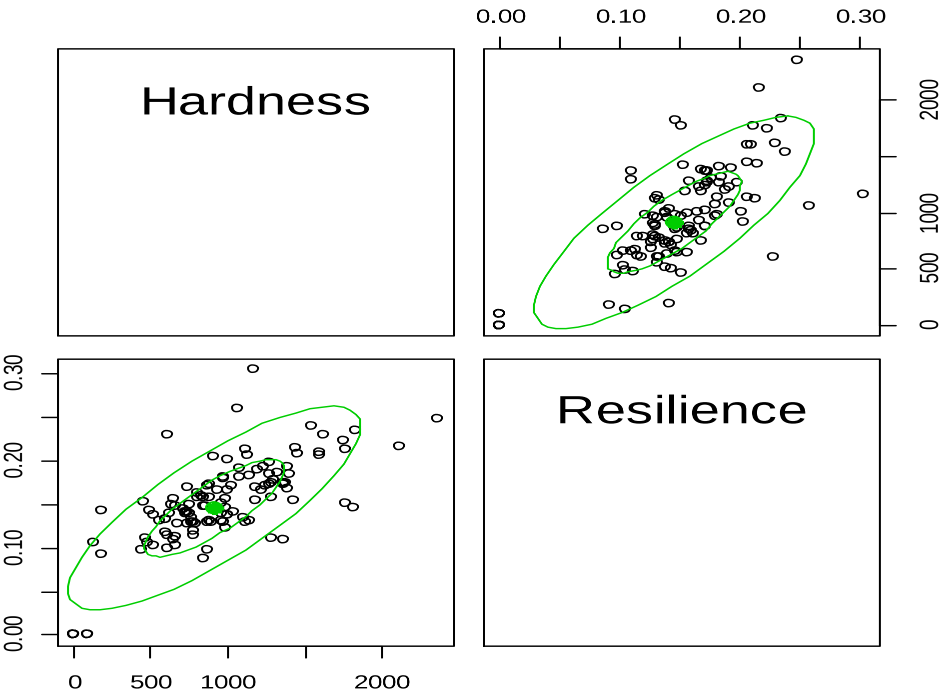


Figure 1: Scatter plot Matrix for the dependent variables

From Figure 2 and Appendix Figure 1, it is clear that the two outcomes variables do not follow the normal distribution. These histogram and density plot though not conditional on the covariates, would possibly give an over view of its distribution. These figures clearly show that the observations are non-negative and right skewed justifying the choice of gamma and lognormal model to describe the positive part of our responses variable.

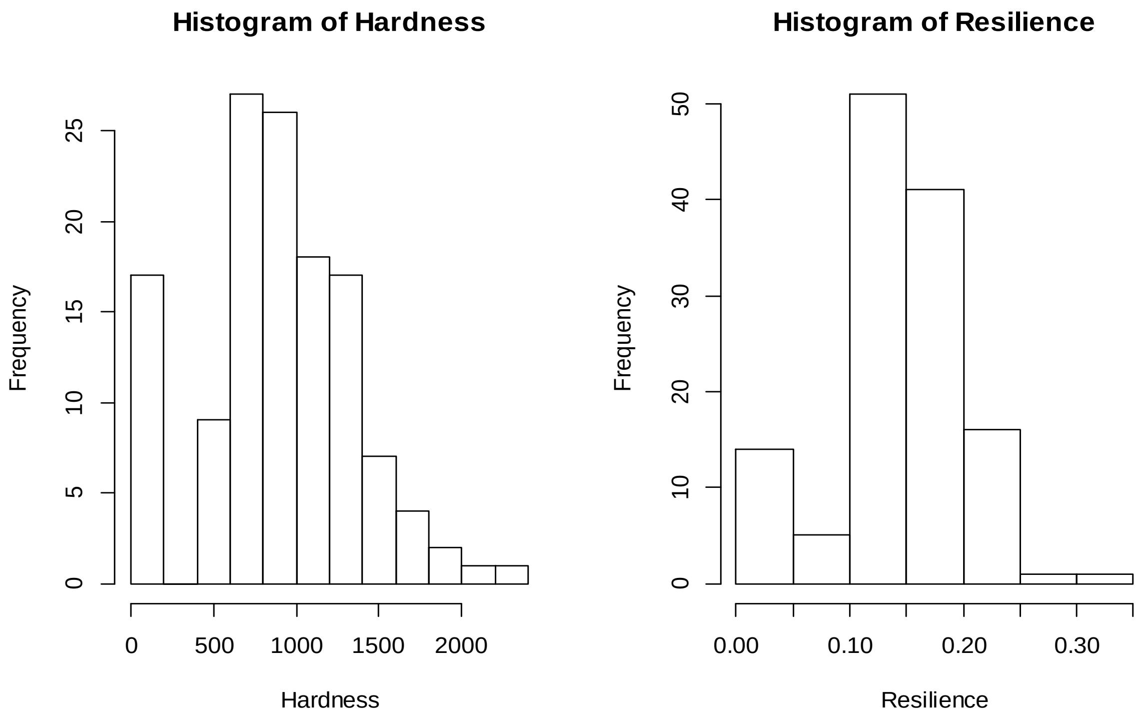


Figure 2: Histograms of the dependent variables

The scatter plot matrix of the dependent variable (Figure 1) revealed a strong correlation between Hardness and Resilience with Pearson correlation computed as 0.8035 with a highly significant p-value ($p < .0001$) (Table 2).

4.2. Statistical Analysis

This research focuses on the estimation and prediction of Hardness and Resilience of meat emulsions for food production. These outcomes herein are of particular interest for manufactures due to the fact that people are prepared to pay a higher price for food or any other manufactured product when the texture is "just right". As mentioned above (Table 1) the data set consists of 129 observations corresponding to different combinations of percentages of seven raw materials and two responses variable. The degree of zeroes values in this sample data set was 10.85% and 7.75% for Resilience and Hardness respectively implying that almost 1/9 and 1/14 of the observations corresponding to different combinations of percentages of the eight raw material contain a end product which was not resilient nor hard enough.

As mentioned by (Amemiya, 1984), “ordinary least squares parameter estimates are biased when applied to the regressions when the outcome contain zeroes values”, we then opted for the zero-inflated gamma (ZIG) and lognormal (ZILN) models with logit link function in the univariate setting. To assess the goodness of fit of the models we compute the AIC and negative log-likelihoods (actually -2* the log-likelihood). A desirable model specification will be the one that fits the data in estimation and is able to predict dependent variable values with accuracy. Table 3 presents all the fitted models while Table 4 and Table 5 present parameter estimates for the ZIG for Resilience outcome and ZILN for Hardness outcome respectively which were found to be the best model for Resilience and Hardness based on the AIC and negative log-likelihoods (the smaller the better). Appendix Table 1 and Appendix Table 2 present the results of the ZILN model for Resilience and ZIG model for Hardness respectively.

In the process of running these ZIG models in SAS, it became apparent that the data did not result in any statistically significant variables used to estimate the rate parameter, $\boldsymbol{\eta}$. Because there were no apparent advantages to estimating this set of variables, regressions were run using only an intercept term for $\boldsymbol{\varphi}$, keeping $\boldsymbol{\eta}$ constant across all observations.

Table 3: Fitted Models

Models	Outcome Variable	Link function	AIC	-2 Log Likelihood
ZILN	Hardness	Logit	**206.9**	**152.9**
	Resilience	logit	-451.1	-509.1
ZIG	Hardness	logit	1797.1	1743.1
	Resilience	logit	**-480.9**	**-538.9**

It can be seen from Table 3 that the ZIG model fits the data better (smaller AIC and -2 Log Likelihood) for Resilience than the ZILN whereas ZILN model fits better than the ZIG for Hardness. Moreover, there is a significant improvement in fit when moving from the ZILN to the ZIG models especially for Hardness. The results indicate that the model fit to our data was improved by using a zero inflated gamma model for Resilience and zero inflated log normal model for Hardness.

4.2.1. Zero-Inflated Gamma Model: Resilience (first outcome variable)

As already mentioned, parameter estimates in zero-inflated model refer to two distinct processes. The first process includes the likelihood of the zero outcomes or the one described by a gamma and/or log normal distribution. This process is estimated through δ utilizing in equation (3). Based on this formulation, the parameter estimates can be expressed as the negative of the marginal impact of the covariate on the probability of a positive outcome, relative to the variance of the Bernoulli component as discussed in the paper by Eric J. Belasco and Sujit K. Ghosh (2008):

$$\delta_k = \frac{\partial \rho_i(\delta)}{\partial x_{ki}} . \frac{1}{\rho_i(\delta)} . \left[\frac{1+\exp(x_i'\delta)}{\exp(x_i'\delta)}\right] = -\frac{\partial \rho_i}{\partial x_{ki}} . \frac{1}{\rho_i(\delta)(1-\rho_i(\delta))} \qquad (6)$$

Where the variance is shown as $\rho_i(\delta)(1-\rho_i(\delta))$. For example from Table 4 and Table 5, except for the intercepts, all the main effects estimates have negative δ coefficient meaning that the raw materials largely and positively influences the likelihood of positive Resilience and/or Hardness of meat emulsion. This is not surprising given that a high percentage of combinations of raw materials tend to contribute to an end product which is harder and more resilient whereas low combinations of percentages of raw materials tend to be more likely to result in Hardness of 0g and/or Resilience of 0 joule/m^3. Therefore, these main effects have a higher probability of incurring positive Resilience and Hardness realizations that can be modeled with the gamma and log normal distributions. These results showed that we did not lose much information by dropping variable 'raw material 8' since the remaining covariates in the model can still contribute to an end product with acceptable texture. It is important to note here that none of the interaction terms was found to be significant for the zero part of the model.

Parameter estimates for γ refer to the marginal impact that the covariates have on the positive realizations of Resilience of meat emulsion. Except of raw materials 1 and 2 which were found to be insignificantly related to Resilience, all the remaining main effects were significantly related to this response variable. Decision was then made to keep the insignificant main effects in the model to enable interpretation of their interaction terms. Due to the presence of significant two and three ways interactions between the covariates (Table 4), the gamma regression coefficients for the variables will no longer indicate a change in the mean gamma of the response with a unit percentage increase of that particular variable alone, keeping the other variables constant; rather it also depends on the level of the other covariates involve in the interaction.

The contribution of variable 'raw material 2' for example to Resilience of meat emulsion after cooking will depend on the contribution of all the other raw materials interacting with raw material 2 (that is raw material 3, 4, 5, 7, interaction between 1 and 5, and between 3 and 6).

Table 4: Estimates standard errors and p-values for the ZIG (Resilience)

Variables	Estimates	Std. Errors	P-Values	Parameters
Probability limit effect (ρ)				
Intercept	3.1439	1.2599	0.0138	
Raw material 1	-6.1314	2.2386	0.007	
Raw material 2	-8.8182	3.03	0.0043	
Raw material 3	-3.4820	1.7152	0.0444	
Raw material 4	-8.6228	2.9854	0.0045	δ
Raw material 5	-11.0164	3.8256	0.0047	
Raw material 6	-19.0017	9.867	0.0563	
Raw material 7	-4.2099	1.8422	0.0239	
Mean effects (μ)				
Intercept	-2.102	0.0654	<.0001	
Raw material 1	-0.122	0.0855	0.1559	
Raw material 2	-0.076	0.1108	0.494	
Raw material 3	0.384	0.1024	0.0003	
Raw material 4	0.5828	0.1014	<.0001	
Raw material 5	0.689	0.0983	<.0001	
Raw material 6	0.2842	0.0893	0.0018	
Raw material 7	0.5981	0.1016	<.0001	
Two ways Interaction				
Raw material 1 × 2	-0.8452	0.3592	0.0201	γ
Raw material 2 × 3	-0.4834	0.4041	0.2338	
Raw material 2 × 4*	-1.6434	0.3791	<.0001	
Raw material 2 × 5	0.7781	0.3755	0.0402	
Raw material 2 × 7	-1.0668	0.379	0.0056	
Raw material 3 × 4	-1.2074	0.3961	0.0028	
Raw material 4 × 5	-0.8809	0.3657	0.0174	
Raw material 4 × 7	-0.7985	0.3742	0.0347	
Raw material 5 × 6	-0.8804	0.3599	0.0158	
Three ways interaction				
Raw material 1 × 2 × 5**	16.1895	3.5457	<.0001	
Raw material 2 × 3 × 4	8.6361	3.7554	0.0231	
Raw material 2 × 3 × 6	-8.3143	3.4725	0.0181	
Rate (η)				φ
Intercept	0.001953	0.000262	<.0001	
Fit Statistics				
-2 Log Likelihood	-538.9			
AIC	-480.9			

*= Interaction between raw material 2 and raw material 4. **= Interaction between raw material 1, 2 and 5.

It can be shown that a change in the mean of gamma with a unit percentage increase in raw material 5 when raw material 2, 4 and 6 and interaction between raw material 1 and 2 is held constant is:

$$\exp\,(0.689 + 0.7781\text{Raw material 2} - 0.8809\text{Raw material 4} \; - 0.8804\text{Raw material 6} + 16.1895\text{Raw material 1} \;\times 2\,)$$

Similarly, the change for example in the mean of gamma for Resilience with a unit percentage increase in raw material 7 when raw material 2 and 4 are held constant is:

$$\exp\,(0.5981 - 1.0668\text{Raw material 2} \; - 0.7985\text{Raw material 4})$$

Hence, in the gamma regression, both the effect of for example raw material 7 for a given level of raw materials 2 and 4 and the effect of raw materials 2 and 4 for given level of raw material 7 depend on the level of the other predictor variables.

Considering the slope of the regression function plotted against raw material 7 now differ for raw materials 2 and 4 = 1% and raw materials 2 and 4 = 2%. The slope of the response (Resilience) function when raw materials 2 and 4 are 1% is given by:

$$\exp(0.5981 - 1.0668 \times 0.01 \; - 0.7985 \times 0.01) = 1.785$$

And when raw materials 2 and 4 = 2%, the slope are:

$$\exp(0.5981 - 1.0668 \times 0.02 \; - 0.7985 \times 0.02) = 1.752$$

Thus, a unit percentage increase in raw materials 2 and 4 has a negative effect on Resilience or decreases the Resilience by 0.033 joule/m^3 (1.785 – 1.752) when the combinations of raw material 2 and 4 are at the higher level than when it is at the lower level. Another way of saying this is that the slopes of the regression lines between Resilience and raw material 7 are different for the different combinations of raw materials 2 and 4. The estimates of the interaction terms between raw material 2 and 7 and raw materials 4 and 7 (-1.0668 and -0.7985) indicate how different those slopes are. This could be due to the interference or antagonistic interactions effect.

4.2.2. Zero-Inflated Log Normal Model: Hardness (Second outcome variable)

The ZILN model was found to be the best fit for the second outcome variable (Hardness). And we can see from Table 5 the estimates for the zero part (δ) of the model are negative except the intercept and can have similar interpretation with those of Table 4.

Parameter estimates for β refer to the marginal impact that the covariates have on the positive realizations of Hardness of meat emulsion. Except of raw materials 1, 2, 5 and 6 which were found to be insignificantly related to Hardness, all the remaining main effects were significantly related to the Hardness of the end product. Decision was then made to keep the insignificant main effects in the model to enable interpretation of their interaction terms between these raw materials; same reason holds for the interactions between raw material 1 and 2 and raw material 2 and 4.

Table 5: Estimates standard errors and p-values for the ZILN (Hardness)

Variables	Estimates	Std. Errors	P-Values	Parameters
Probability limit effect (ρ)				
Intercept	5.0892	1.7779	0.0049	
Raw material 1	-7.2320	2.6487	0.0072	
Raw material 2	-10.7094	3.8409	0.0061	
Raw material 3	-10.1332	3.5836	0.0054	(δ)
Raw material 4	-10.1864	3.6116	0.0056	
Raw material 5	-13.3559	5.0333	0.009	
Raw material 6	-22.5224	10.5389	0.0345	
Raw material 7	-13.2705	4.9917	0.0088	
Mean effects (μ)				
Intercept	6.7600	0.2178	<.0001	
Raw material 1	-0.5868	0.3227	0.0713	
Raw material 2	0.4228	0.3034	0.1658	
Raw material 3	-1.249	0.3169	0.0001	
Raw material 4	0.5935	0.2877	0.0412	
Raw material 5	0.2987	0.3037	0.3271	
Raw material 6	-0.1938	0.313	0.5369	
Raw material 7	-0.8776	0.3316	0.0091	
Two ways Interaction				
Raw material 1 × 2	-0.9693	1.2016	0.4214	
Raw material 1 × 3	5.4846	1.2703	<.0001	(β)
Raw material 1 × 7	4.5682	1.2352	0.0003	
Raw material 2 × 4	-1.7365	1.2045	0.1518	
Raw material 3 × 5	3.8274	1.1952	0.0017	
Raw material 3 × 6	4.0011	1.1583	0.0007	
Raw material 5 × 7	4.4628	1.1625	0.0002	
Raw material 6 × 7	3.408	1.1651	0.0041	
Three ways interaction				
Raw material 1 × 2 × 3	-41.0594	11.9081	0.0008	
Raw material 2 × 4 × 7	-43.1015	11.5839	0.0003	
Variance (α)				
intercept	-0.8957	0.06482	<.0001	(α)
Fit Statistics				
-2 Log Likelihood	152.9			
AIC	206.9			

The contribution of raw material 5 to Hardness of meat emulsion will depend on the presence of raw material 3 and raw material 7 as can be seen from Table 5, there is a significant interaction effect between the three raw materials.

For example, considering the slope of the regression function plotted against raw material 5 now differ for raw materials 3 and 7 = 1% and raw materials 3 and 7 = 2%. The slope of the response (Hardness) function when raw materials 2 and 4 are 1% is given by:

$$(0.2987 + 3.8274 \times 0.01 + 4.4628 \times 0.01) = 0.381$$

And when raw materials 3 and 7 = 2%, the slope are:

$$(0.2987 + 3.8274 \times 0.02 + 4.4628 \times 0.02) = 0.464$$

Thus, a unit percentage increase in raw materials 3 and 7 has a slightly positive effect on Hardness or increases the Hardness by 0.083g (0.464 – 0.381) in the logarithmic scale when the combinations of raw material 3 and 7 are at the higher level than when it is at the lower level. Another way of saying this is that the slopes of the regression lines between Hardness and raw material 5 are different for the different combinations of raw materials 3 and 7. The estimates of the interaction terms between raw material 3 and 5 and raw materials 7 and 5 (3.8274 and4.4628) indicate how different those slopes are. This could be due to the interference or antagonistic interactions effect.

These outcomes were also modeled using a zero-inflated log normal (ZILN) and ZIG models. Tough the AIC opted for ZIG model for Resilience outcome, the parameter estimates for this model were found to be similar to those of ZILN for this particular outcome. In the other hand ZILN model for Hardness showed also similar output with variable ‘Raw Material 7’ becoming significant whereas this was not the case with ZIG model for Hardness. These models characterize the positive observations using a log normal and gamma distributions, which can take into account highly skewed data. The results from the ZILN model for Resilience and ZIG model for Hardness are shown in Appendix Table 1 and Appendix Table 2 respectively.

5. Discussion and Conclusion

It was the purpose in this project to understand the effect of raw materials on Hardness and Resilience of meat emulsions for food production, understand how it is possible to improve these characteristics of the final product after cooking by changing the percentages of the raw materials. Data for 129 observations corresponding to different combinations of percentages of raw materials with Hardness and Resilience as outcomes variables was analyzed. From the EDA it was clear that the data was positively skewed with outcomes containing zeroes values. And we would not therefore expect the classical regression models to fit the data well (Amemiya, 1984).

Modeling data sets with the presence of zeroes values in the outcome variables remains a large problem in statistics. While use of the Tobit model may be well-justified in certain instances such as the cattle feeder example discusses in the paper by Eric J. Belasco and Sujit K. Ghosh (2008), the results from meat emulsions data sets suggest the use of a zero-inflated modeling mechanism. This is particularly true in instances where data come from a two-step process. While two-step processes have been applied to hurdle models, zero-inflated models have largely been ignored in profitable studies. This is mainly a result of the past limitation of zero-inflated models to count data.

In this project, a zero-inflated model is developed that can handle both univariate and multivariate situations rather efficiently. Additionally, the inherent parametric flexibility allows for distributional assumptions to change based on the data on hand, rather than strictly using normally distributions. Here we use a log-normal and gamma distributions to capture the positively skewed nature of Hardness and Resilience of meat emulsion after cooking, which gives the zero-inflated model significant advantages over the Tobit model. Though it is argued that the log-normal and the gamma density function are both likely to give equally fitting results (Wiens, 1999), Kundu and Manglicky (2005) extend from Wiens contention (Wiens, 1999) that there are cases when the interchangeability of these two models slightly differs. Particularly on tail data which is crucial on inferences. It was therefore worthwhile fitting both of them and examining the result. The ZIG model seemed to give a better fit for Resilience outcome variables with smaller AIC values than the ZILN model whereas ZILN fitted better than ZIG for Hardness outcome variable. Advantages in model fit for the ZILN and ZIG model stem from the ability of the zero-inflated model to isolate the impacts from observing a positive Hardness and Resilience and their levels.

It emerged from Bernoulli part of ZIG and ZILN models that all the main effect estimates have negative δ coefficient meaning that the raw materials largely and positively influences the likelihood of positive Resilience and/or Hardness of meat emulsion. From the gamma and log normal part, raw materials 3, 4, 5, 6 and 7 were found to be significantly related to Resilience while raw materials 3, 4 and 7 were significantly related to Hardness of meat emulsion. These results are in agreement with those obtained from the raw data. As expected both in the gamma and lognormal part of the models, the interactions terms (two and three ways) were found to be significant on both outcomes. These presences of significant interactions indicate that the effect of one raw material on Hardness and/or Resilience is different at different values of the other raw materials.

Knowledge of Hardness and Resilience of processing meat emulsion is very crucial for production companies to ensuing product acceptability. Results from this research demonstrate the potential gains from using this particular mixture model. Additionally, the ZIG demonstrated a strong ability to fit the data slightly better than the ZILN for Resilience whereas for Hardness the ZILN was better than ZIG model.

In conclusion, had they been there were no significant interactions terms in the model, improvement of Hardness and/or Resilience of the end product after cooking could have been made by combining different combinations of percentages of the significant mean effects; that is combining different combinations of percentages of raw materials 3, 4, 5, 6 and 7 for Resilience and raw materials 3, 4 and 7 for Hardness. The presence of significant two and three ways interaction terms in the models drastically changes the interpretation of all of the coefficients (main effects) and becomes more challenging since the gamma regression coefficients for example raw materials 2 and 4 will no longer indicate the relationship between these raw materials and Resilience keeping the other variables constant because of the interaction between them; rather it also depends on the level of the other variable involve in the interaction. It is also important to mention here that though raw material 5 was not significantly contributing to Hardness of meat emulsion, this was due to the presence of interactions terms in the model and particular attention should be paid to this raw material as can be seen from Appendix Table 3, variable 'raw material 5' was found to be highly significantly related to both outcomes. Moreover, privilege should be given to the estimates with high estimates because we suspect that their different combinations of percentages could lead to an end product which is acceptably hard and resilient.

6. Limitations and recommendations

Often, multivariate tests are more powerful especially when the responses are highly correlated. Regression from univariate Hardness and Resilience models offer information concerning the relative impacts each covariate has on these outcomes variable. However, these variables would be likely better characterized in multivariate setting in order to capture the covariance structure between Hardness and Resilience; one of the advantages of doing multivariate analyses is that you can conduct tests of the coefficients across the different models. This was attempted in this research but due to the complexity of the model and time constraints, the model did not sufficiently converge. We therefore recommend future research to consider the multivariate setting and possibly examine other possible characteristics that might explain Hardness and Resilience of meat emulsion, or other possible covariates related to these characteristics of the end product after cooking in order to ensure product acceptability.

References

Akaike, H. (1974) 'A new look at the statistical model identification.' *IEEE Transactions on Automatic Control* 19, 716–723.

Amemiya, T. (1974) 'Multivariate regression and simultaneous equation models when the dependent variables are truncated normal.' *Econometrica* 42(6), 999–1012.

Amemiya, (1984). 'Tobit models: A survey.' *Journal of Econometrics* 24, 3–61.

Arabmazar, A., and P. Schmidt (1982) 'An investigation of the robustness of the tobit estimator to non-normality.' *Econometrica* 50(4), 1055–1064.

Belasco, E. J., B. K. Goodwin, and S. K. Ghosh (2007) 'A multivariate evaluation of ex-ante risks associated with fed cattle production.' Selected Paper. SCC-76: Economics and Management of Risk in Agriculture and Natural Resources Annual Meeting.

Belasco, E. J., M. R. Taylor, B. K. Goodwin, and T. C. Schroeder (2006) 'Probabilistic models of yield, price, and revenue risk for fed cattle production.' Selected Paper. American Agricultural Economics Association Annual Meeting, Long Beach.

Bera, A. K., C. M. Jarque, and L. F. Lee (1984) 'Testing the normality assumption in limited dependent variable models.' *International Economic Review* 25(3), 563–578.

Burnham, K. P. , and Anderson, D. R., (2002). *Model selection and multimodel inference: A practical information-theoretic approach*, third edition. New York: Springer-Verlag.

Chavas, J. P., and K. Kim (2004) 'A heteroskedastic multivariate tobit analysis of price dynamics in the presence of price floors.' *American Journal of Agricultural Economics* 86(3), 576–593.

Cornick, J., T. L. Cox, and B.W. Gould (1994) 'Fluid milk purchase: A multivariate tobit analysis.' *American Journal of Agricultural Economics* 76, 74–82.

Critical Reviews in Food Science and Nutrition, 38(3):173–258 (1998) – *journal.*

Eiswerth, M. E., and J. S. Shonkwiler (2006) 'Examining post-wildfire reseeding on arid rangeland: A multivariate tobit modeling approach.' *Econological Modeling* 192, 286–298

Eric J. Belasco and Sujit K. Ghosh (2008) '*Modeling Censored Data Using Mixture Regression Models with an Application to Cattle Production Yields*'.

FOOD ENGINEERING – Vol. II – *Texture in Solid and Semisolid Foods* – Bourne, Malcolm.

Jones, A. M. (1989) 'A double-hurdle model of cigarette consumption.' *Journal of Applied Econometrics* 4(1), 23–39.

Keele, Luke. (2008). *Semiparametric regression for the social sciences.* West Sussex, England: John Wiley & Sons Ltd

Kundu, D. & Manglicky A., *Discriminating Between The Log-normal and Gamma Distributions* Journal of Statistical Planning and Inference Volume 127, Issues 1-2, 1 January 2005, Pages 213-227.

Lee, L. F. (1993) 'Multivariate tobit models in econometrics.' *In Handbook of Statistics*, Vol. 11, ed. G. S. Maddala, C. R. Rao, and H. D. Vinod. chapter 6, pp. 145–173.

Lesiów, T., and Xiong, Y. L. (2003). Chicken muscle homogenate gelation properties: effect of pH and muscle fiber type. *Meat Science*, 64, 399-403.

Mathew Flynn, Ph.D. Louise A., Francis FCAS, MAAA (winter 2009) '*More Flexible GLMs: Zero-Inflated Models and Hybrid Models*'.

Terrell, R. N. (1983). Reducing the sodium content of processed meats. *Food Technology*, 37,66-71.

Verbeke G. and Molenberghs G.: *Project: Longitudinal Data Analysis*. Uhasselt; 2008/2009.

Wiens, B.L. (1999), When log-normal and gamma models give different results: a case study, The American Statistician, 53, 2, 89-93.

http://www.freepatentsonline.com/EP0265740.html (last accessed 18/11/2012)

Appendix

Zero-Inflated Log Normal model

Resilience

Appendix Table 1: Estimates, standard errors and p-values for ZILN (Resilience)

Variables	Estimates	Std. Errors	P-Values	Parameters
Probability limit effect (ρ)				
Intercept	3.1357	1.2585	0.014	
Raw material 1	-6.1208	2.2366	0.0071	
Raw material 2	-8.8026	3.0263	0.0043	
Raw material 3	-3.4738	1.714	0.0447	
Raw material 4	-8.6061	2.981	0.0046	(δ)
Raw material 5	-10.9925	3.8183	0.0047	
Raw material 6	-19.0111	9.8985	0.057	
Raw material 7	-4.201	1.8408	0.0241	
Mean effects (μ)				
Intercept	0.1182	0.0115	<.0001	
Raw material 1	-0.0174	0.0150	0.2489	
Raw material 2	-0.0079	0.0195	0.6822	
Raw material 3	0.0593	0.0179	0.0012	
Raw material 4	0.0972	0.0178	<.0001	
Raw material 5	0.1152	0.0169	<.0001	
Raw material 6	0.0445	0.0156	0.0053	
Raw material 7	0.0965	0.0178	<.0001	
Two ways Interaction				
Raw material 1 × 2	-0.07204	0.0635	0.2588	(β)
Raw material 2 × 3	-0.0735	0.0712	0.3042	
Raw material 2 × 4*	-0.2459	0.0667	0.0003	
Raw material 2 × 5	0.1372	0.0663	0.0405	
Raw material 2 × 7	-0.1794	0.0663	0.0078	
Raw material 3 × 4	-0.2037	0.0689	0.0037	
Raw material 4 × 5	-0.1659	0.0636	0.0103	
Raw material 4 × 7	-0.1363	0.0654	0.0392	
Raw material 5 × 6	-0.1552	0.0630	0.0151	
Three ways interaction				
Raw material 1 × 2 × 5	2.7977	0.6217	<.0001	
Raw material 2 × 3 × 4	1.2433	0.6515	0.0586	
Raw material 2 × 3 × 6	-1.0436	0.6071	0.088	

Variance (α)				
intercept	-3.844	0.06596	<.0001	(α)
Fit Statistics				
-2 Log Likelihood	-509.1			
AIC	-451.1			

Zero Inflated Gamma: Hardness

Hardness (ZIG)

Appendix Table 2: Estimates, standard errors and p-values for ZIG (Hardness)

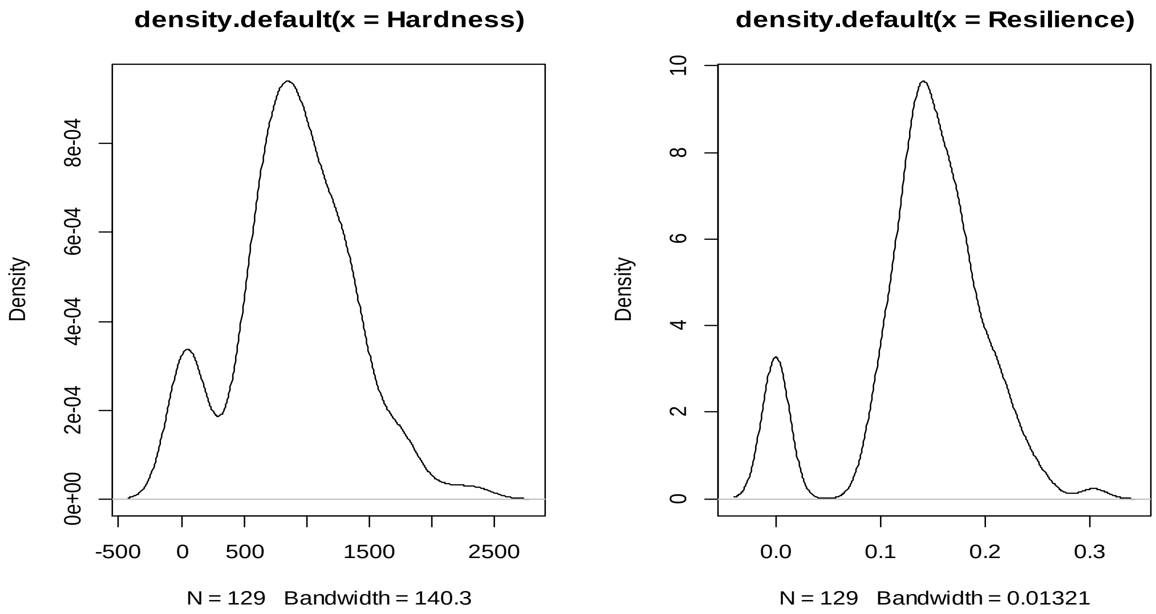
Variables	Estimates	Std. Errors	P-Values	Parameters
Probability limit effect (ρ)				
Intercept	5.0861	1.7769	0.0049	
Raw material 1	-7.2279	2.6472	0.0072	
Raw material 2	-10.7051	3.8394	0.0061	
Raw material 3	-10.1579	3.5935	0.0055	
Raw material 4	-10.1526	3.5987	0.0055	(δ)
Raw material 5	-13.3075	5.0113	0.0089	
Raw material 6	-22.481	10.499	0.0341	
Raw material 7	-13.3093	5.0099	0.0089	
Mean effects (μ)				
Intercept	6.7231	0.211	<.0001	
Raw material 1	-0.5447	0.3071	0.0784	
Raw material 2	0.3649	0.2762	0.1888	
Raw material 3	-0.8489	0.3192	0.0088	
Raw material 4	0.6336	0.2656	0.0185	
Raw material 5	0.3896	0.2901	0.1817	
Raw material 6	-0.0508	0.2901	0.8612	
Raw material 7	-0.4192	0.3371	0.216	
Two ways interaction				
Raw material 1 × 2	-0.6405	1.0895	0.5576	
Raw material 1 × 3	4.7701	1.1995	0.0001	
Raw material 1 × 7	3.6454	1.1508	0.0019	(γ)
Raw material 2 × 4	-1.4249	1.1093	0.2013	
Raw material 3 × 5	3.1592	1.1464	0.0067	
Raw material 3 × 6	3.0396	1.0743	0.0054	
Raw material 5 × 7	3.609	1.1027	0.0014	
Raw material 6 × 7	3.1883	1.0504	0.0029	
Three ways interaction				
Raw material 1 × 2 × 3	-42.069	11.2089	0.0003	
Raw material 2 × 4 × 7	-47.1066	11.1151	<.0001	

Rate (η)				**(φ)**
Intercept	106.04	16.0182	<.0001	
Fit Statistics				
-2 Log Likelihood	1743.1			
AIC	1797.1			

Appendix Table 3: Correlation Matrix between the responses and covariates

Pearson Correlation Coefficients, N = 129 Prob > |r|	R_M1	R_M2	R_M3	R_M4	R_M5	R_M6	R_M7	R_M8	Hardness	Resilience
R_M1	1.00000	-0.1414	-0.1465	-0.1449	-0.1449	-0.1415	-0.1480	-0.1474	-0.12261	-0.13586
		0.1098	0.0976	0.1012	0.1012	0.1096	0.0941	0.0955	0.1663	0.1248
R_M2	-0.1414	1.00000	-0.1430	-0.1373	-0.1415	-0.1421	-0.1446	-0.1391	-0.04673	-0.13834
	0.1098		0.1057	0.1206	0.1096	0.1081	0.1020	0.1138	0.5990	0.1179
R_M3	-0.1465	-0.1430	1.00000	-0.1425	-0.1425	-0.1389	-0.1414	-0.1449	-0.08160	-0.10147
	0.0976	0.1057		0.1072	0.1072	0.1164	0.1098	0.1011	0.3579	0.2525
R_M4	-0.1449	-0.1373	-0.1425	1.00000	-0.1409	-0.1415	-0.1440	-0.1434	0.10727	0.16514
	0.1012	0.1206	0.1072		0.1111	0.1096	0.1034	0.1048	0.2263	0.0615
R_M5*	-0.1449	-0.1415	-0.1425	-0.1409	1.00000	-0.1415	-0.1398	-0.1434	0.34916	0.41754
	0.1012	0.1096	0.1072	0.1111		0.1096	0.1138	0.1048	<.0001	<.0001
R_M6	-0.1415	-0.1421	-0.1389	-0.1415	-0.1415	1.00000	-0.1404	-0.1440	0.10966	0.15485
	0.1096	0.1081	0.1164	0.1096	0.1096		0.1123	0.1034	0.2161	0.0797
R_M7	-0.1480	-0.1446	-0.1414	-0.1440	-0.1398	-0.1404	1.00000	-0.1465	0.12985	0.05850
	0.0941	0.1020	0.1098	0.1034	0.1138	0.1123		0.0976	0.1425	0.5102
R_M8	-0.1474	-0.1399	-0.1449	-0.1434	-0.1434	-0.1440	-0.1465	1.00000	-0.43760	-0.41253
	0.0955	0.1138	0.1011	0.1048	0.1048	0.1034	0.0976		<.0001	<.0001
Hardness	-0.1226	-0.0467	-0.0816	0.10727	0.34916	0.10966	0.12985	-0.4376	1.00000	0.80354
	0.1663	0.5990	0.3579	0.2263	<.0001	0.2161	0.1425	<.0001		<.0001
Resilience	-0.1358	-0.1383	-0.1014	0.16514	0.41754	0.15485	0.05850	-0.4125	0.80354	1.00000
	0.1248	0.1179	0.2525	0.0615	<.0001	0.0797	0.5102	<.0001	<.0001	

R_M5* = Raw Material 5



Appendix Figure 1: Density plot of the response variables

Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
Understanding the effect of raw materials on some characteristics of meat emulsions for pet food production

Richting: **Master of Statistics-Epidemiology & Public Health Methodology**
Jaar: **2013**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt
behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -,
vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten
verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Voor akkoord,

Nango, Cabrale

Datum: **4/02/2013**