

ESTIMATING MELAMINE POPULATION DISTRIBUTIONS WITH LEFT-CENSORED DATA

Hsien-Tang Tsai, Department of Business Management, National Sun Yat-sen University, Kaohsiung, TAIWAN, htt@mail.nsysu.edu.tw

Shih-Hsiung Lo, Department of Business Management, National Sun Yat-sen University, Kaohsiung, TAIWAN, lo2330@gmail.com

Tung-Ju Wu, Department of Business Management, National Sun Yat-sen University, Kaohsiung, TAIWAN, n_bear007@yahoo.com.tw

ABSTRACT

The event of contaminated food by melamine had drawn worldwide attentions concerning food safety issues. The aim of this article is to analyze the test results on melamine collected from the website of the Center for Food Safety of Hong Kong. Since the values of detected melamine level below the legal limit (LOD) are not reported and the products are classified as “satisfactory”, the data are left-censored which can be analyzed by using Turnbull estimator and the results can serve as criteria for comparison among various product categories. Three product categories are analyzed under the assumption of lognormal distribution for comparisons, and the results can provide good references for decision makers.

INTRODUCTION

In 2007, wheat gluten, protein sources, pet foods and animal feed in North America were found to be contaminated with melamine and related compounds, such as cyanuric acid, ammeline and ammelide [1]. Numerous pet foods in the United States were recalled after dogs and cats consuming the products suffered renal failure. In 2008, high levels of melamine were detected in infant formula and other liquid and powdered milk products originating from China. These high levels of melamine in infant milk and other milk products have led to severe health effects and illness in Chinese infants and young children [1]. In 2008, The Food and Drug Administration (FDA) found melamine contaminants had been added to vegetable proteins imported from China and had been used as ingredients in pet food [2]. FDA also discovered infant formula contaminated with melamine and cyanuric acid made by China manufacturer [1].

According to Review of “WHO Expert consultation to the toxicological aspects of melamine and cyanuric acid” [3], 294,000 infants have been found to be affected by the melamine-contaminated infant formula and have been treated for renal complications as of the end of November 2008 in China. The report went on to add that more than 50,000 infants have been hospitalized, and six children have died as a direct result of the melamine contamination. World Health Organization and international food safety and health authorities strictly banned foods melamine contamination in 2009 [4].

The Hong Kong government introduced legislation with a view to prohibiting inappropriate level of melamine in food sold in Hong Kong in September 2008. Actions had been taken out by the Centre for Food Safety (CFS) to monitor melamine level in food sold in Hong Kong at all levels of the food trade. Testing were continued for dairy products and extended to cover all products containing milk elements imported from countries other than mainland China as far as practicable. Full information on testing results was posted on its web site [5].

In this article, we first collect and compare the test results on melamine, posted by the Center for Food

Safety of Hong Kong on its website, for difference product categories. Second, since the values of detected melamine level below the legal limit (LOD) are not reported and the product is classified as “satisfactory”, the data are left-censored which can be analyzed by using Turnbull estimator and the results can serve as criteria for comparison among various product categories.

METHODS FOR SAMPLES WITH NON-DETECTS

Turnbull [6] proposes a self-consistency algorithm for parametric estimation of the distribution function for arbitrary censoring. For randomly left censored data, the PLE (product limit estimator) is defined as follows. Let $a_1 < a_2 < \dots < a_M$ be the M distinct values at which detects occur, r_j is the number of detects at a_j , and n_j is the sum of non-detects and detects that are less than or equal to a_j . Then the PLE is defined to be 0 for $0 \leq x \leq a^*$, where a^* is a_1 or the value of the detection limit for the smallest non-detect if it is less than a_1 . For $a^* \leq x < a_M$ the PLE is $F_j = j \prod (n_j - r_j) / n_j$, where the product is over all $a_j > x$, and the PLE is 1 for $x \geq a_M$. When there are only detects this reduces to the usual definition of the empirical cumulative distribution function (Frome and Wambach, [7]). The maximum likelihood estimation (MLE) method with various desirable properties such as consistency, asymptotic unbiasedness, and efficiency is a useful statistical method to fit parametric models and obtain estimates that maximize the likelihood of the empirical censored data (Jin, et al. [8]).

Let Y denote the observed value of melamine contained in food from test. Since its measurement scale is based on ppm (10^{-6}), it is reasonable to assume that Y is log-normal distributed. Y follows a lognormal distribution with geometric mean (GM) e^μ and geometric standard deviation (GSD) e^σ . The probability density function of y is given by

$$f(y | \mu, \sigma) = e^{-(\log y - \mu)^2 / 2\sigma^2} / \sqrt{2\pi}\sigma y, \quad y > 0 \quad (1)$$

The probability of observing y less than the limit of detection (LOD) is given by

$$P(y < LOD) = \Phi\left(\frac{\log(LOD) - \mu}{\sigma}\right) \quad (2)$$

where $\log(\cdot)$ denotes the natural logarithm and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution with mean zero and standard deviation one (Jin, et al. [8]). Let Y denote the observed value of melamine contained in food from test sample, then Y is log-normal distributed, denoted by $Y \sim \text{LN}(\mu_y, \sigma_y^2)$. A transformed new variable $Z = \ln Y$ is a normal random variable with a mean μ_z and a variance σ_z^2 , denoted by $Z \sim N(\mu_z, \sigma_z^2)$. Theoretically, if Y is $\text{LN}(\mu_y, \sigma_y^2)$ then the mean and variance of Y are $E(Y) = \mu_y = \exp(\mu_z + \sigma_z^2 / 2)$, $\text{Var}(Y) = \sigma_y^2 = \exp(2\mu_z + \sigma_z^2)[\exp(\sigma_z^2) - 1]$.

The Proc LIFEREG in SAS uses an iterative algorithm developed by Turnbull [6] to compute a parametric maximum likelihood estimate of the cumulative distribution function for left, right, and interval censored data, and create probability plots for the data. (SAS Institute, [9]). Therefore, in this article, we adopted the Turnbull estimation methods to estimate parametric statistics for left-censored empirical data of melamine under the lognormal assumption.

DATA SOURCE

The Centre for Food Safety (CFS), the Government of the Hong Kong Special Administrative Region, released the “test results on melamine” during September 2008 to April 2009 on its website (www.cfs.gov.hk [5]), after then, test results on melamine will be incorporated to CFS food safety report under Food Surveillance Programme. The CFS adopts a risk-based approach in collecting and analyzing

food samples for melamine. With the approach, the CFS can execute a systematic and comprehensive surveillance of products likely to be affected by melamine. The CFS completed one round of sampling for all food products commonly available in the local market. The product categories consist of infant formula milk powder (below 12 months), baby food, milk powder, milk and milk beverage (other than infant formula milk powder), frozen confections (ice-cream, ice bar, yoghurt, etc), other milk products (butter, cheese, cream, evaporated milk, etc), and other food product (chocolate / candies, cookies / biscuits / egg rolls, Chinese bun / cake / bread, nutritional supplements, Eggs, poultry meat, meat, fish and aquatics products, vegetable, miscellaneous). The information of test results consists of product name, manufacturer's name, date of sampling and testing result in ppm.

The Government has established legal limits for melamine in food since late September 2008, as follow: 1 mg/kg for milk and food for children under the age of 36 months, pregnant and lactating women, and at 2.5 mg/kg for other food. Using the legal limits as the benchmarks, products with melamine levels exceeding the legal limits were classified as unsatisfactory products, while products with levels complying with the legal limits were classified as satisfactory products. The legal limit for other food product is 2.5 mg/kg and 1 mg/kg for the others. Note that products contain melamine level between 1 to 2.5 mg/kg were determined satisfactory.

RESULTS

The Centre for Food Safety (CFS) of Hong Kong published test results on melamine during 2008/09-2009/04 on its website(www.cfs.gov.hk), 4649 samples were collected in total, which include infant formula milk powder (below 12 months), baby food, milk powder, milk and milk beverage (other than infant formula milk powder), frozen confections (ice-cream, ice bar, yoghurt, etc), other milk product (butter, cheese, cream, evaporated milk, etc), and other food product (chocolate / candies, cookies / biscuits / egg rolls, Chinese bun / cake / bread, nutritional supplements, Eggs, poultry meat, meat, fish and aquatics products, vegetable, miscellaneous). Among them, there were only 39 unsatisfactory samples, and the overall satisfactory rate was 99.992%.

Table 1 lists numbers of satisfactory, numbers of unsatisfactory, and unsatisfactory rate for each product category. There are eight categories with zero rates. The higher unsatisfactory rates occur in other food product, "Chinese bun/ cake / bread" (0.021) is highest; "cookies, biscuits, eggrolls" (0.019) is second; "chocolate /candies" (0.014), eggs (0.014) are third. Ammonium bicarbonate is used in the food industry as a raising agent for flat baked goods, such as cookies, crackers, steamed buns and Chinese almond cookies. Since ammonium bicarbonate from China has been found to be tainted with melamine, it leads to the high unsatisfactory rate of biscuits, Chinese bun/ cake / bread.

Three product categories of milk & milk beverages, frozen confections (ice-cream, ice-bar, yoghurt, etc), and other food products (chocolate/candies, cookies, biscuits, eggrolls, eggs, Chinese bun/cake/bread) are analyzed by using the Turnbull estimator in Proc Lifereg in SAS to estimate the population parameters based on the left-censored data under the lognormal assumption. Since the detected melamine level of a satisfactory product is not reported, i.e. the values below the legal limit are censored, and a LOD for each category is set as its legal limit and shown in Table 2. Since the intercept and scale denotes the mean (μ) and standard deviation (σ) for non-censored distribution under the lognormal assumption in (1), they provide criteria of comparison for various product categories and both are the-small-the-better. The intercept of frozen confections (-15.8674) is the smallest, however, the scale (6.4652) is the largest. If scales are around the same levels (3.7486 vs 3.1435), then the larger intercept (-5.7446) of other food products deserves more attention for correction.

TABLE 1: Test results on melamine from Hong Kong during 2008/09-2009/04.

Product Category	Satisfactory	Unsatisfactory	Unsatisfactory rate
Infant Formula Milk Powder (<12month)	156	0	0.000
Baby Food	251	1	0.004
Milk Powder (Other than Infant Formula Milk Powder)	151	0	0.000
Milk and Milk Beverage	810	5	0.006
Frozen Confections (ice-cream, ice-bar, yoghurt, etc)	710	5	0.007
Other Milk Product (butter, cheese, cream, evaporated milk, etc)	194	0	0.000
Other Food Products			
-Chocolate/Candies	359	5	0.014
-Cookies, biscuits, eggrolls	616	12	0.019
-Eggs	290	4	0.014
-Chinese bun / cake / bread	236	5	0.021
-Nutritional supplements	76	0	0.000
-Poultry meat	72	0	0.000
-Meat (Pork / Beef / Mutton)	62	0	0.000
-Fish & Aquatic Products	141	0	0.000
-Vegetables	23	0	0.000
-Miscellaneous	463	2	0.004
Total	4610	39	0.008

TABLE 2: Summary parameter estimates for left-censored melamine data from Hong Kong.

Methods	Total observations	Number of non-censored	LOD (ppm)	Intercept	Scale	GM	GSD
Milk and Milk Beverages	815	5	1	-9.3793	3.7486	0	42.46
Frozen Confections	710	5	1	-15.8674	6.4652	0	642.36
Other Food Products	1527	26	5	-5.7446	3.1435	0.0032	23.1842

CONCLUSIONS

Through the lesson of contaminated food by melamine, the issue of food safety had deserved worldwide attentions once more. The Center for Food Safety of Hong Kong released all test results on melamine during September 2008 to April 2009 on its website, which can effectively prevent the occurrences of contaminated events. Since the values of detected melamine below the legal limit (LOD) are not reported and the product is classified as “satisfactory”, the data are left-censored which can be analyzed by using Turnbull estimator in Proc Lifetest of SAS and the results can serve as criteria for comparison among various categories. Three product categories are analyzed under the assumption of lognormal distribution for comparisons, and the results can provide good references for decision makers.

ACKNOWLEDGEMENT

This research is supported by National Science Council, R.O.C under the grant NSC 98-2410-H-110 -024 -MY3 to National Sun Yat-sen University.

REFERENCES

- [1] Hilts, C., & Pelletier, L. Background paper on occurrence of melamine in foods and feed”, http://www.who.int/foodsafety/fs_management/Melamine_3.pdf, [accessed 1 Apr 2010].
- [2] Sun, H., Wang, L., & Ge, X. The research advancement for analytical methodology of melamine and related analogues in environment and foods. *Chemical Journal on Internet*, 2009, 11(8), 37-55.
- [3] WHO. “Experts set tolerable level for melamine intake”. <http://www.who.int/mediacentre/news/releases/2008/pr48/en/> [accessed 14 Apr 2010], 2008.
- [4] Freeonline Library. The melamine incident: implications for international food and feed safety”. <http://www.thefreelibrary.com/The+melamine+incident:+implications+for+international+food+and+feed...-a0220845104> [accessed 14 Apr 2010], 2009.
- [5] Centre for Food Safety, the Government of the Hong Kong Special Administrative Region. Test results on melamine. http://www.cfs.gov.hk/english/whatsnew/whatsnew_fstr/melamine_result.html (accessed 2011/10/23).
- [6] Turnbull, B.W. The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society, Series B (Methodological)*, 1976, 38(3), 290-295.
- [7] Frome, E.L., & Wambach, P.F. Statistical methods and software for the analysis of occupational exposure data with non-detectable values. Prepared by Oak Ridge National Laboratory for the US Department of Energy, 2005.
- [8] Jin, Y., Hein, M.J., Deddens, J.A., & Hines, C.J. Analysis of lognormally distributed exposure data with repeated measures and values below the limit of detection using SAS. *Ann. Occup. Hyg.*, 2011, 55(1), 97-112.
- [9] SAS Institute. *SAS/STAT® 9.2 User’s Guide the LIFEREG procedure (Book Excerpt)*, Cary, NC: SAS Institute Inc., 2009.