Food Sci. Biotechnol. Vol. 17, No. 3, pp. 659 ~ 663 (2008)

Food Science and Biotechnology

The Korean Society of Food Science and Technology

RESEARCH NOTE

Distribution Functions Describing the Microbiological Contamination of Seasoned Soybean Sprouts

Jin Pyo Park[1], Dong Sun Lee[2], and Hyun-Dong Paik*

Division of Animal Life Science, Konkuk University, Seoul 143-701, Korea

[1]Department of Computer Engineering, Kyungnam University, Masan, Gyeongnam 631-701, Korea

[2]Department of Food Science and Biotechnology, Kyungnam University, Masan, Gyeongnam 631-701, Korea

Abstract Different statistical distribution functions were examined to find an adequate distribution function to describe the microbial contamination behavior of a Korean side dish product, seasoned soybean sprouts for different seasons and market groups. The triang distribution was the best for any market groups in winter, while the logistic distribution could describe the microbial contamination in log CFU/g for all the market groups in spring and summer. From parametric bootstrapping based on the fitted distributions, it was found that a normal distribution could describe the distribution of mean microbial count in log CFU/g for all the seasons and market groups. Statistical parameters for each season/market group are presented to estimate the confidence interval.

Keywords: microbial count, statistical distribution, Korean side dish, bootstrap confidence interval

Introduction

Quality assurance programs regarding Korean side dishes prepared with mixing dry spices and chill-marketed should be based on microbial criteria because of their vulnerability to microbial spoilage and the high correlation between microbial quality and sensory properties (1,2). However, microbial count data used for quality control are usually display irregular and fluctuating patterns. The description of microbial count group by distribution functions can be very useful to estimate the general behavior of microbial contamination and the frequency of incidence of high microbial outbursts outside the tolerable range (3-5). Corradini *et al.* (6) tried to estimate the frequency of microbial counts in commercial food products by using various distribution functions.

The distribution function for microbial contamination can be useful for assessing the exposure level of pathogenic and spoilage microorganisms (7-9). A correct description of microbial contamination distribution can also help to design logistic control and shelf life determinations for the food (10). However, there is very limited information on the distribution functions of microbial contamination of prepared foods. To our knowledge there are no reports on the use of statistical distribution function to represent microbial contamination of Korean seasoned side dishes, which are increasing in prepared food sales in Korea.

Therefore this study aims to identify adequate distribution function to describe the microbial contamination behavior of a Korean side dish product, seasoned soybean sprouts.

Materials and Methods

As a source data set of the microbial contamination of Korean side dish products, aerobic bacterial counts reported by Park *et al.* (11) for seasoned soybean sprouts of different seasons and market groups were used for statistical analysis. Forty microbial counts expressed as log CFU/g for each season/market group were plotted and analyzed using the @RISK Program (Palisade Corporation, Ithaca, NY, USA), which fits the distributions to the data using the maximum likelihood estimation (MLE) method to find an adequate distribution. The distributions were selected based on the ranking according to the *p*-value of chi-squared statistics, and then a quantile-quantile (Q-Q) plot was also employed to ascertain goodness-of-fit (12).

As a further step for estimating the microbial distribution and confidence interval for the mean microbial counts, the parametric bootstrapping method was applied based on the fitted distributions as described above (13). The 1,000 bootstrap samples were generated for each season/market group combination and then fitted again with a proper distribution function, which was also confirmed by Q-Q plot (12).

Results and Discussion

Among the many candidate distributions examined for describing the variability of microbial counts for each season/market group, logistic and triang distributions were the most relevant based on the *p*-values (Table 1). Variation of aerobic bacterial counts in log CFU/g for any market group in winter season could be represented by the triang distribution function as given by:

$$f(x)=\frac{2(\max-x)}{(\text{m.likely}-\min)(\max-\min)},\ \text{for } \min<x<\text{m.likely} \quad (1)$$

$$f(x)=\frac{2(\max-x)}{(\min-\text{m.likely})(\max-\min)},\ \text{for } \text{m.likely}<x<\max \quad (2)$$

where f(x) is the probability distribution function as a

---

*Corresponding author: Tel: +82-55-2049-6011; Fax: +82-55-455-3082
E-mail: hdpaik@konkuk.ac.kr

Received August 29, 2007; Revised October 27, 2007; Accepted October 28, 2007

Table 1. The *p*-values of chi-square statistics of the fitted distribution, and correlation coefficients between its quantile and quantile of experimental microbial counts

Seasons/Market type	Adopted distribution function	*p*-value	Correlation coefficient (r)
Winter/Traditional market	Triang	0.934	0.998
Spring/Traditional market	Logistic	0.342	0.946
Summer/Traditional market	Logistic	0.587	0.988
Winter/Discount store	Triang	0.996	0.997
Spring/Discount store	Logistic	0.167	0.872
Summer/Discount store	Logistic	0.824	0.769
Winter/Department store	Triang	0.450	0.994
Spring/Department store	Logistic	0.824	0.992
Summer/Department store	Logistic	0.369	0.742

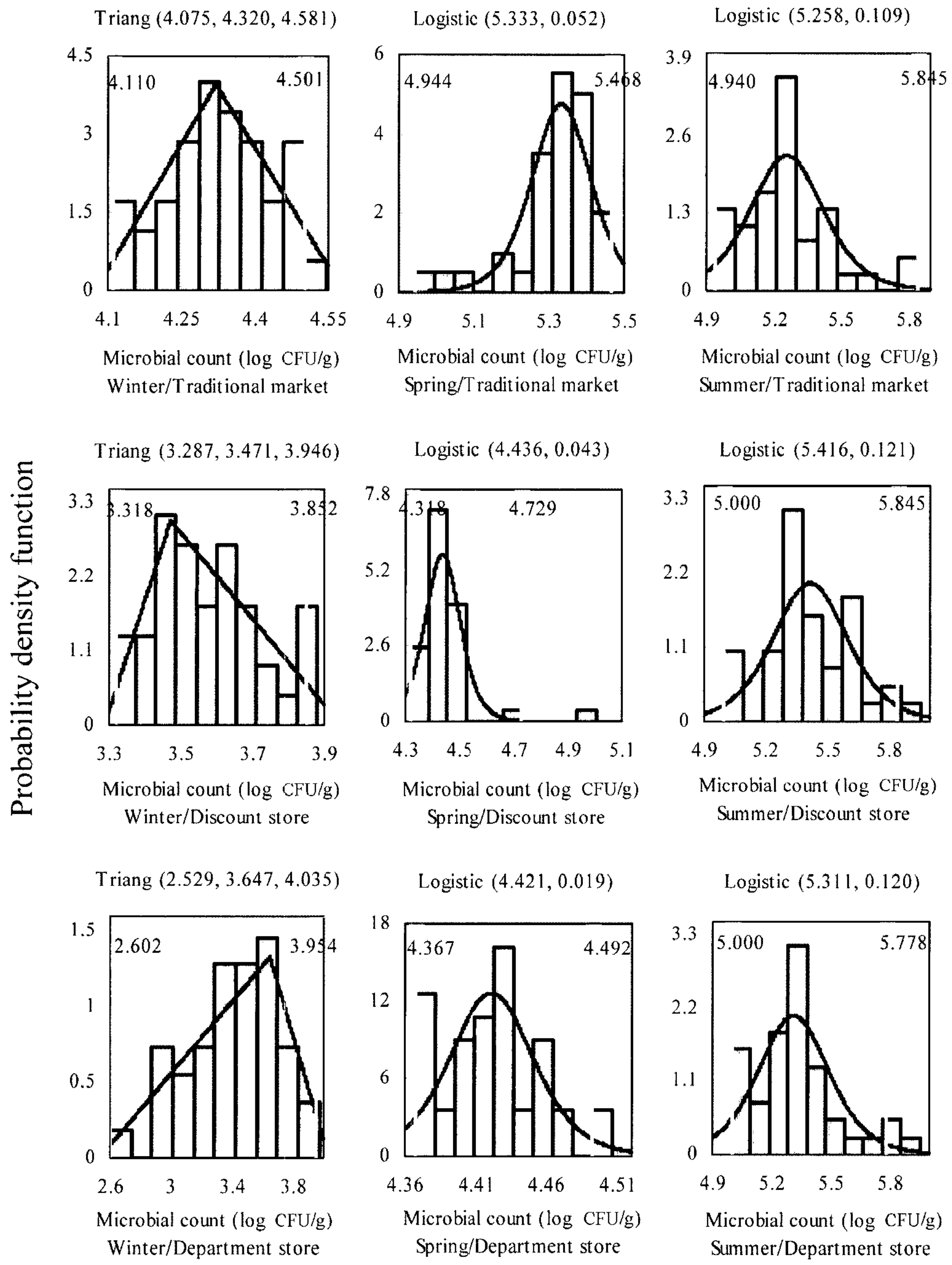

Fig. 1. Comparison of fitted probability density functions and histograms of various microbial counts of soybean sprouts. Two vertical lines indicate a 95% confidence interval of the microbial count. Estimations of each distribution parameters are given in the bracket of each function.

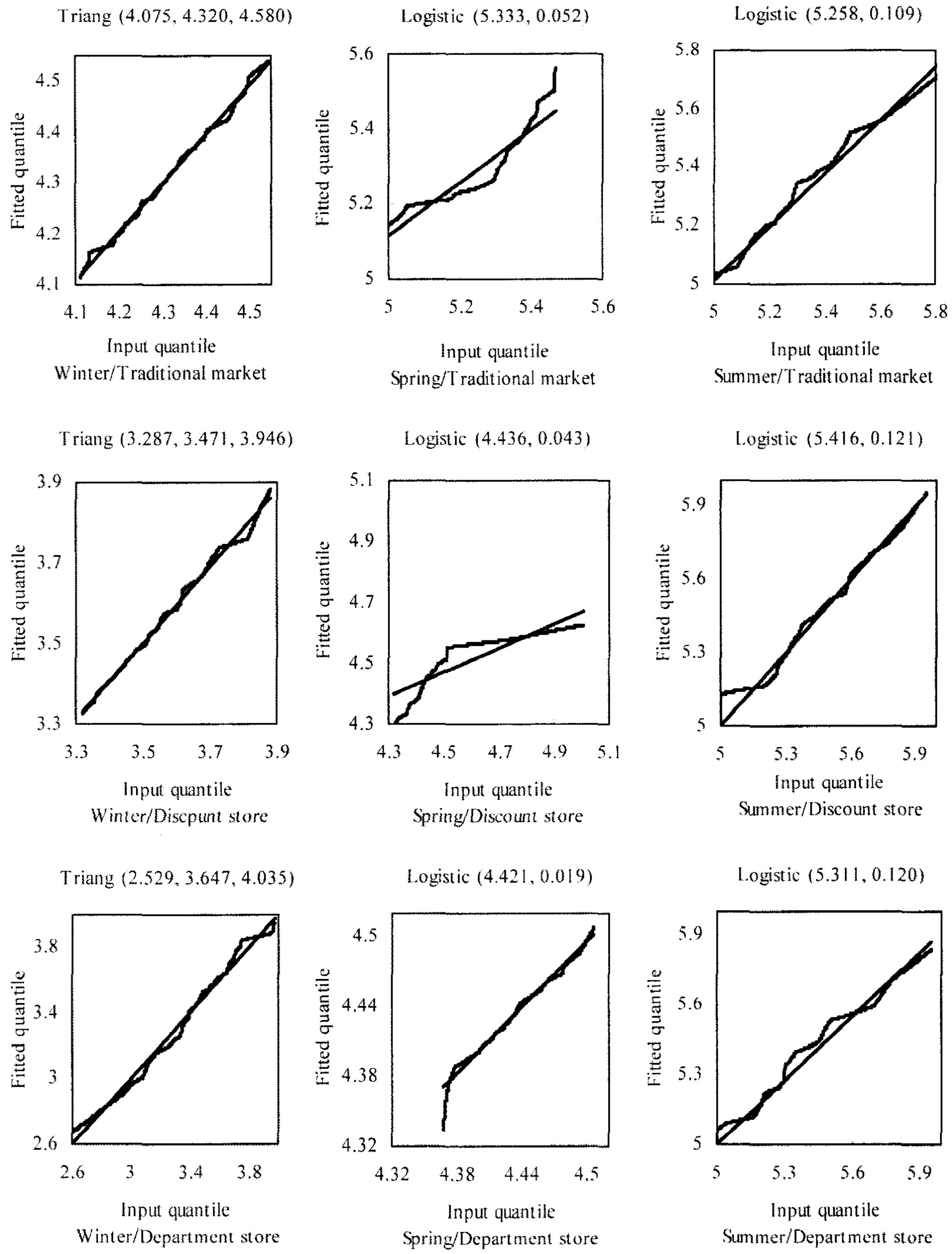

Fig. 2. Quantile-quantile (Q-Q) plots for the distribution of Fig. 1.

function of the microbial count (log CFU/g), min is minimum value, m.likely is continuous mode parameter of the most likely value, and max is maximum value.

For the seasons of spring and summer, the most suitable is the logistic distribution function written as:

$$f(x)=\frac{\exp\left(-\frac{x-\alpha}{\beta}\right)}{\beta\left\{1+\exp\left(-\frac{x-\alpha}{\beta}\right)\right\}^2}, \quad \text{for } -\infty<x<\infty \tag{3}$$

where α is continuous location parameter and β is continuous scale parameter other than 0.

Plots in Fig. 1 overlay histogram plots of the observed microbial counts with probability density functions of fitted distributions. These plots offer a visual comparison between the shape of the data and the fitted distributions.

The goodness of the fit to the distribution model was evaluated by examining the p-values of chi-square statistics (Table 1) and Q-Q plots (quantile-quantile comparison between data distribution and fitted distribution, Fig. 2): the better the fit, the larger the p-value is; the better the fit, the closer the Q-Q plot resembles an overlaid straight line. The correlation coefficients between the quantile of fitted distribution and the quantile of microbial counts were also calculated to investigate the linearity of the Q-Q plot (Table 1): the closer the correlation coefficient is to 1, the stronger

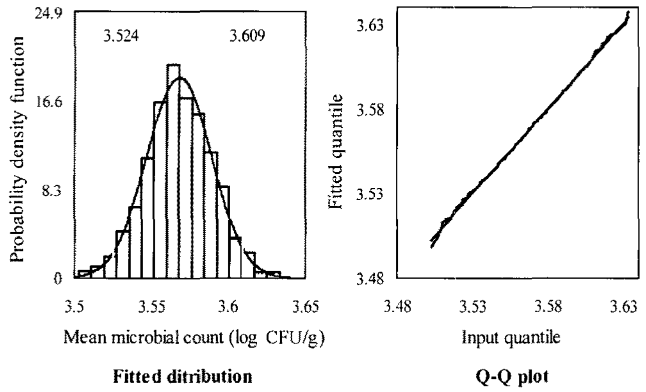

Fig. 3. The bootstrap distribution and Q-Q plot for mean microbial counts of the winter/discount store group according to a normal distribution (3.568, 0.021). Two vertical lines indicate 95% confidence interval of the microbial count.

the linearity is. When examined according to these criteria, traditional market and discount stores in the winter season showed the best goodness of the fit to the triang distribution, while discount stores in the spring were the worst with logistic distribution. Examining Q-Q plots also reveals that fitness to a logistic distribution of any market group in the spring season was worse than that of other seasons. It is noted that the spring season involves greater fluctuations in temperature through processing and marketing, which may cause irregularity in the microbial quality of the product. Considering that relatively small sample sizes were used for this statistical analysis, the irregularity and abruptness in the microbial quality may have disturbed the randomness of the data sampled. In the future, improved goodness of the fit may be obtained if a larger amount of the data is collected.

Based on the present analysis, the distribution functions given in Table 1 and Fig. 1 and 2, are generally the best options available from the available functions: triang distribution for the market groups in winter and logistic distribution for the market groups in spring and summer. Corradini *et al.* (6) reported that no single distribution was consistently outstanding among those examined (normal, log normal, Laplace, log Laplace, Weibull, extreme value, beta, and log beta) to estimate the high microbial count occurrence in various foods of raw milk, dairy products, and frozen apple juice, even though normal distribution worked fairly well in many cases. The initial contamination of aerobic bacteria, anaerobic bacteria, thermophiles, and *Bacillus cereus* has been described by several distributions of log normal, beta and log Laplace distributions (5,7-9, 14). The probability density functions of the fitted distributions with estimated parameters presented in Fig. 1 would be useful to characterize the variability of microbial counts for any season/market group: the probability that a microbial count will be equal or greater than any given value, the moment of distribution can be calculated and used for quality assurance programs (6); a range of shelf life estimation can be given rather than a single value (9, 10). According to Peleg *et al.* (5), successive lots of food products in a factory could be described by similar loads of the same distribution function, which underlines the usefulness of the distribution characterization.

In order to estimate the mean contamination level for any season/market type, parametric bootstrapping based on the fitted distributions was conducted and a typical result is shown in Fig. 3. A normal distribution could describe the distribution of mean microbial count in log CFU/g for all the seasons and market groups as shown by respective Q-Q plots. The outcome is consistent with the central limit theorem that the distribution of means even from non-normal distribution samples generally follows the shape of a normal distribution with a relatively large sample size (12,13). The fitted distribution is helpful to execute statistical inference for the population mean. Therefore, the normal distribution given in Table 2 can be used to estimate the mean microbial count for any season and market group. 95% bootstrap confidence intervals for the mean contamination level of seasoned soybean sprouts consistent with a normal distribution are also provided in Table 2.

Acknowledgments

This work was supported by Korea Science and Engineering Foundation (Project # R01-2005-000-10235-0) and Brain Korea 21 program, Korea.

References

1. Kim G-T, Ko Y-D, Lee DS. Shelf life determination of Korean seasoned side dishes. Food Sci. Technol. Int. 9: 257-263 (2003)
2. Seo I, Park JP, Lee DS. Correlation between microbiological and sensory quality indices of Korean seasoned side dishes stored under chilled conditions. J. Food Sci. Nutr. 11: 257-260 (2006)
3. Engel R, Normand MD, Horowitz J, Peleg M. A qualitative

Table 2. The 95% bootstrap confidence intervals and adapted distribution for mean contamination levels of seasoned soybean sprouts as a function of season and market group

Season/Market group	95% Bootstrap confidence interval	Adapted distribution
Winter/Traditional market	4.294-4.357	Normal (4.325, 0.016)
Spring/Traditional market	5.294-5.370	Normal (5.332, 0.019)
Summer/Traditional market	5.199-5.317	Normal (5.258, 0.030)
Winter/Discount store	3.524-3.601	Normal (3.568, 0.021)
Spring/Discount store	4.411-4.459	Normal (4.436, 0.012)
Summer/Discount store	5.348-5.482	Normal (5.415, 0.035)
Winter/Department store	3.305-3.499	Normal (3.403, 0.051)
Spring/Department store	4.410-4.431	Normal (4.421, 0.005)
Summer/Department store	5.242-5.377	Normal (5.310, 0.034)

probabilistic model of microbial outbursts in foods. J. Sci. Food Agr. 81: 1250-1262 (2001)

4. Gonzalez-Martinez C, Corradini MG, Peleg M. Probabilistic models of food microbial safety and nutritional quality. J. Food Eng. 56: 135-142 (2003)
5. Peleg M, Nussinovitch A, Horowitz J. Intrepretation of and extraction of useful information from irregular fluctuating industrial microbial counts. J. Food Sci. 65: 740-746 (2000)
6. Corradini MG, Normand MD, Nussinovitch A, Horowitz J, Peleg M. Estimating the frequency of high microbial counts in commercial food products using various distribution functions. J. Food Protect. 64: 674-681 (2001)
7. Delignette-Muller ML, Rosso L. Biological variability and exposure assessment. Int. J. Food Microbiol. 58: 203-212 (2000)
8. Bahk G-J, Todd ECD, Hong C-H, Oh D-H, Ha S-D. Exposure assessment for *Bacillus cereus* in ready-to-eat *kimbab* selling at stores. Food Control 18: 682-688 (2007)
9. Lee DS, Hwang K-J, Seo I, Park JP, Paik H-D. Estimation of shelf life distribution of seasoned soybean sprouts using the probability of *Bacillus cereus* contamination and growth. Food Sci. Biotechnol. 15: 773-777 (2006)
10. Giannakourou MC, Koutsoumanis K, Nychas GJE, Taoukis PS. Development and assessment of an intelligent shelf life decision system for quality optimization of the food chill chain. J. Food Protect. 64: 1051-1057 (2001)
11. Park JP, Kim HW, Lee DS, Paik H-D. Seasonal and market group variation in the microbiological quality of seasoned soybean sprouts. Food Sci. Biotechnol. 16: 325-328 (2007)
12. Scheaffer RL, McClave JT. Probability and Statistics for Engineers. Duxbury Press, Belmont, CA, USA. pp. 222-455 (1995)
13. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. Chapman & Hall/CRC, Boca Raton, FL, USA. pp. 39-59 (1993)
14. Peleg M. Intrepretation of the irregular fluctuating microbial counts in commercial dairy products. Int. Dairy J. 12: 255-262 (2002)