J. Fd Hyg. Safety 18(4), 237-242 (2003)

식품산업체에서 확률분포 모델을 이용한 불규칙적인 미생물 수 분포 활용

박경진† · 김성조 · 심우창 · 천석조 · 최원상* · 홍종해**

*한국보건산업진흥원 HACCP팀, *동국대학교 자연과학대학 생명공학과, **강원대학교 수의학과*

Application of the Fluctuating Microbial Counts Using Probability Approaches in Food Industries

Gyung-Jin Bahk†, Sung-Jo Kim, Woo-Chang Sim, Seok-Jo Chun, Weon Sang Choi* and Chong-Hae Hong**

HACCP Team, Korea Health Industry Development Institute, Seoul 156-050, Korea

**Department of Biotechnology, College of Natural Sciences, Dongguk University, Gyeongju 780-714, Korea*

***Department of Veterinary Medicine, Kangwon National University, Chunchon, Kangwon-Do 200-701, Korea*

ABSTRACT – Sequences of industrial microbial counts of foods shows irregular fluctuating patterns as a definition of fluctuating microbial counts(FMC). Recently, it became clear that the FMC was considered as having a lognormal distribution as a first order approximation. Application of lognormal distribution to the industrial microbial counts could produce useful information in practice. This study is intended to verify the application method of lognormal distribution in FMC. The one year's records of microbial counts of frozen dumplings from two companies were obtained, and the statistical analysis was carried out to estimate the frequencies of future events where counts exceeding selected levels and to compare the sanitation level of the two companies. The results showed that this application method enable translation of irregular records of microbial counts into an useful information such as the actual probabilities of outburst of a given level and the quantitative predictions of potential hazards in the processing.

Key words: Fluctuating microbial counts, lognormal distribution, frozen dumplings

식품산업체에서 일반세균 수 등의 미생물 분석결과는 품질평가 기준으로 이용될 뿐만 아니라, 등급분류와 제품의 안전성 판단 기준으로도 많이 이용된다. 하지만 정량적인 미생물 분석결과는 미생물 수의 불규칙 변화(Fluctuating microbial counts, FMC)에 따른 해석에 어려움이 있어 이러한 변화가 나타내는 어떤 특성 혹은 인과관계를 규명하지 못하고 있다. 따라서 그 동안 대부분의 정량적 미생물 자료는 설정된 기준과의 비교에 한정되어 주로 활용되고 있다[1].

미생물 수의 변화는 작업환경내의 여러 요인에 영향을 받는데, 이러한 조건들이 일정하지 않아 동일한 원료로부터 생산된 제품도 롯트 별 또는 일자별로 다양하게 나타난다. 또한 식품에 존재하는 미생물 집락의 특성, 크기 또는 초기 균수 등 미생물 자체의 특성과도 관련되는 것으로 알려져 있다[2]. 미생물 수가 특별히 높게 나타나는 경우는 작업장내 시설, 장비, 기기 등의 개선이나 엄격한 위생관리에 의해 해결되는 경우도 있지만, 알려지지 않은 요인에 의해서 복합적으로 나타나는 경우는 원인 제거가 어렵다. 따라서 후자의 경우 원인 규명을 위한 다른 접근방법이 가능하다면 식품의 안전성 관리수준을 한단계 높이는데 절대적인 도움이 될 것이다.

최근 불규칙적이며 주기적인 발생 양상을 나타내지 않는 일반세균 수에 대해서 무작위적(random) 분포에 기초한 확률적 모델을 활용하여 해석하는 방법론이 제안되었고[3,4]. 이러한 방법론은 Nussinovitch[5], Peleg[6]에 의해서 실제로 몇몇 식품을 통해 증명되었다. 확률적 모델에 의한 FMC 활용을 위해서는 불규칙적이며, 무작위적으로 독립적이고, 특별한 경향을 갖지 않으며, 일반세균 수 발생에 영향을 미칠 수 있는 공정내 일반위생 및 기타 다른 요인들의 변화가 없어야 한다는 조건과 부합되어야 한다. 만약 이러한 조건과 일치한다면 일반세균 수의 분포는 근사값으로서 Lognormal 분포로 추정될 수 있으며, 더 나아가 어떤 특정 기준값(threshold)을 초과하는 확률을 산출하는데 이용할 수 있다는 것이다[4].

본 연구에서는 식품산업체에서 자체검사하고 있는 일반세균 오염실태 자료를 확률분포 모델에 적용함으로써 현장에서의 활용 가능성을 확인하고자 하였다.

†Author to whom correspondence should be addressed.

기본 개념과 모델

FMC를 통하여 미래의 발생가능성을 추정하기 위해서는 식품산업체에서 수집된 일반세균 수 등의 분포자료가 사전에 다음과 같은 가정에 일치하여야 한다[5,6]. 즉,

① 각 롯트의 자료는 독립적으로 수행된 것으로 무작위적(random)이어야 한다.

② 전체적으로 어떤 특정한 경향을 가지고 있지 말아야 한다.

③ 일정기간 수집된 자료의 일반세균 수는 대수적 평균과 표준편차를 가지고 있어야 하며, 일상적인 위생조건에서 분석된 자료로, 극도의 오염 발생이나 위생적인 개선과 같은 특징적인 발생 분포가 없어야 한다.

위의 가정에 만족하는 결과를 표현할 수 있는 가장 간단한 모델 중의 하나는 다음 (1) 식의 Lognormal 분포에 기초하는 것으로 알려져 있다[3,4].

$$N(n) = 10^{\mu L + \sigma L Z(n)} \quad (1)$$

N(*n*) : the count of the nth lot
μL : 10 logarithmic mean of the counts
σL : 10 logarithmic standard deviation of the counts
Z(*n*) : a random number whose magnitude has a standard normal distribution(i.e., μ=0, σ=1)

식(1)은 수학적으로 단순하고 확률을 추정하는데 있어 쉽다는 장점이 있다. Weibull, Beta 등의 확률분포모델도 이용가능하지만 이들을 현재 일반적으로 이용될 수 있는 자료보다도 더 많은 자료를 요구하므로 식품산업체의 응용에 있어서 편리성을 고려하여 Lognormal 분포가 권장되고 있다[4]. 일반세균수의 분포가 독립적으로 수행된 것으로 Lognormal 분포를 따르고 대수적 평균과 표준편차를 가지고 있다면 어떤 특정수준(Nc)을 초과할 확률P(N(*n*)≥Nc)의 산출도 가능하며, 이 계산은 일반적인 소프트웨어인 Excel 등을 이용하여 산출할 수 있다.

재료 및 방법

자료수집

일반세균수 자료는 냉동만두류를 생산하는 2개의 식품업체로부터 각각 1년치의 자료를 수집하였다. 이 자료는 어떤 특정 연구나 실험을 목적으로 위해서 수행된 것이 아닌 일상적으로 위생관리를 위하여 조사하고 분석된 자료로 한정하였으며, 실제 위생적인 조건에 대한 정보는 없었다. 또한 총 분석횟수 등은 한정하지 않았고 Lot수에 대한 요구도 하지 않았다. 즉, 자료는 선택된 자료가 아니며, 순수한 raw data를 수집하여 사전에 스크리닝이나 분류는 하지 않고 Lot 순서로만 정리하였다. 수집된 자료에 사전에 알고 있는 극도의 오염 발생이나 위생적인 개선과 같은 특징적인 분포가 없음을 확인하였다.

자료처리

각각의 count는 대수값(로그값)을 전환하고 이들 2개 업체의 자료들은 각각 전체를 1/2씩 나누어 첫 번째 1/2에서 대수평균(μ_L)과 표준편차(σ_L)를 산출하였으며, 나머지 1/2은 특정수준(Nc)를 초과하는 확률을 검증하는데 이용하였다. 즉, 확률적으로 추정된 수와 실제 나머지 1/2에서 발생한 수를 비교하여 검증에 이용하였다.

자료분석

처리된 자료는 독립성 검정을 위하여 자기상관(Auto-correlation)분석을 실시하였고, Lognormality test를 위하여 Quantile-Quantile plot를 실시하였다. 자기상관 분석은 Peleg[7]에 의한 Excel 에서 사용할 수 있는 Macro 프로그램인 Microbial Count Probability를 이용하였다. 자기상관분석은 여러 다른 Lag에 의해 분리된 연속값들 사이의 linear correlation의 정도를 나타내는 것으로 불규칙적인 연속에서 주기(periodicity)의 존재여부를 결정하기 위하여 사용되었다. Lognormality test 역시 Peleg[7]에 의한 Excel 에서 사용할 수 있는 Macro 프로그램인 Microbial Count Probability를 이용하여 Quantile-Quantile plot(Q-Q plot)에 의하여 상관계수(correlation coefficient)가 1에 근접하는 정도와 Q-Q plot가 직선에 가까운지 정도에 의해 결정하였다.

또한 수집된 자료가 독립성과 Lognormality가 있다는 가정하에 대수적 평균(μ_L)과 표준편차(σ_L)를 이용, 다음 식(2)에 의거 미래에 특정수준(Nc)을 초과할 확률 P(N(*n*)≥Nc)을 계산하였다.

$$N(n) \geq Nc = P(N(n) \geq Nc) * n_{2ndhalf} \quad (2)$$

$n_{2ndhalf}$ = The total number of counts in the second half of the record

결 과

자료의 독립성 여부

2개 업체의 냉동만두제품으로부터 일반세균수의 기록들은 Fig. 1~2와 같이 매우 불규칙적이며 톱니모양을 나타내고 있다. 특정 Lot에서 일시적인 증가 현상이 보이나 이러한 발생

에 대한 특별한 관련요인은 없었고, 표면적으로도 어떠한 경향이나 주기를 가지고 있지 않은 것으로 보였다. 이는 Fig. 3~4의 자기상관 분석에서와 같이 모두 random series로 영점(zrero)을 중심으로 진동하며 특정 모양을 인식할 수 있는 패턴이 없고, 모든 lag에서의 자기상관계수(Autocorrelation coffefficient)가 A 업체의 경우 95% 신뢰수준인 0.438, B 업체의 경우 0.283 보다 낮으므로 모두 유의하지 않았다. 따라서 A, B 업체의 일반세균수 분석은 독립적으로 수행되었다고 볼 수 있다(Table 1).

자료의 Lognormality 여부

Fig. 5~6에서와 같이 Q-Q plot가 직선에 가깝고 Table 1의 상관계수가 A 업체는 0.972, B 업체는 0.946으로 1에 가까우며, 또한 Fig. 7~8과 같이 막대그래프의 분포가 적어도 표면적으로는 lognormal 분포를 가지는 자료로부터 온 것과 같은 모양 즉, 한쪽으로 약간 치우친 형태를 가지고 있으므로, 2개 업체의 자료는 lognormal 분포를 한다고 볼 수 있다.

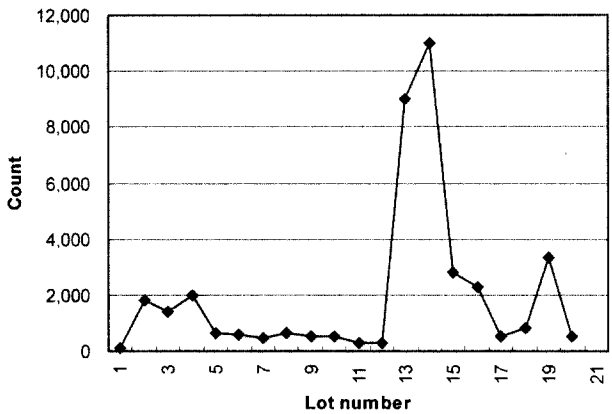

Fig. 1. Total microbial counts of frozen dumplings from source A.

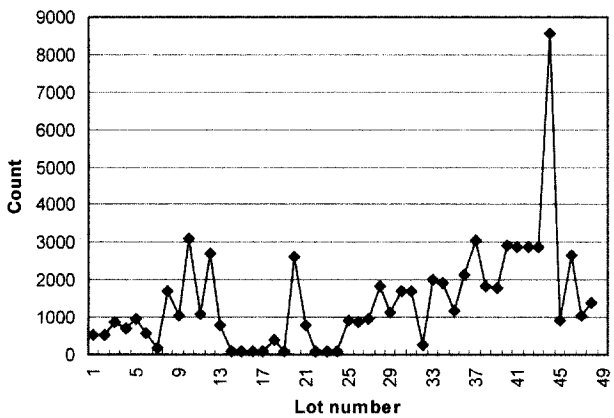

Fig. 2. Total microbial counts of frozen dumplings from source B.

모델에 의한 추정값 검정

첫 번째 1/2을 이용하여 산출된 대수적 평균과 표준편차를 이용하여 식(1)에서 산출한 특정수준(Nc)을 초과하는 확률값과 식(2)에 의한 나머지 1/2에서 Nc를 초과하는 추정값은 Table 1과 같다. 또한 Table 1은 실제 나머지 1/2에서의 Nc를 초과하는 수를 관찰치로 나타내어 비교검증 하였는데, 결과적으로 비록 초기의 모델이지만 작은 데이터베이스에도 불구하고 추정치와 관찰치는 어느 정도 일치하는 것으로 파악되었다. 특히 Nc가 높을수록 일치하는 정도는 더욱 크다고 볼 수 있었다. 그리고 Table 2에는 일반세균수 발생에 있어서의 최소와 최대규모와 대수적 평균과 표준편차를 나타내었는데 특정 수준 Nc (10^5 CFU/g)에 대한 초과할 확률값은 A 업체가 0.298, B 업체가 0.296로 B업체가 상대적으로 양호한 위생수준을 갖고 있는 것으로 정량적으로 추정할 수 있다. 이

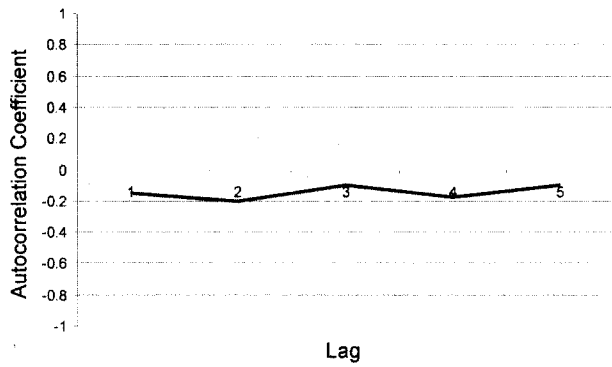

Fig. 3. The autocorrelation functions of the series of the total microbial counts of frozen dumplings from source A.

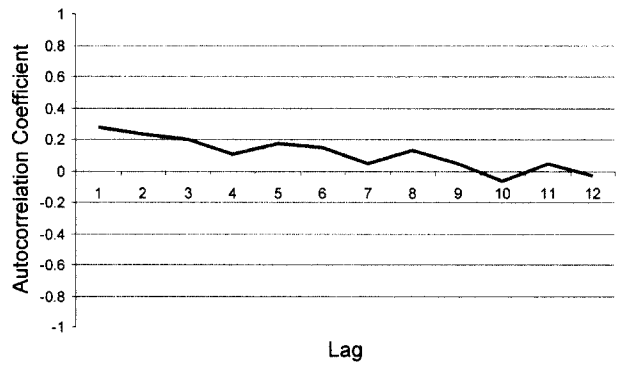

Fig. 4. The autocorrelation functions of the series of the total microbial counts of frozen dumplings from source B.

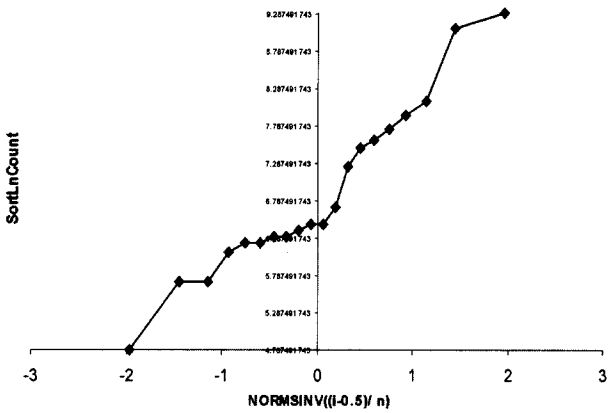

Fig. 5. Lognormal probability (Quantile-Quantile) plot of the series of the total microbial counts of frozen dumplings from source A.

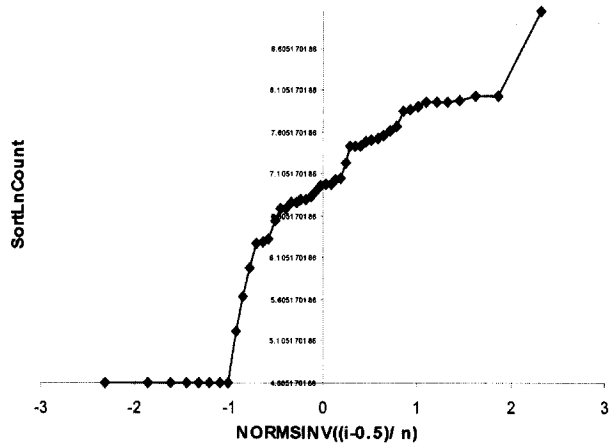

Fig. 6. Lognormal probability (Quantile-Quantile) plot of the series of the total microbial counts of frozen dumplings from source B.

러한 결과 해석은 과거 최소값과 최대값, 대수적 평균과 표준편차, Fig. 1~2와 같은 자료에서는 쉽게 얻을 수 없는 것이다.

고 찰

일반세균 수의 불규칙적인 분포는 어떤 특별한 의미를 부여할 수 없어 일명 "noise" 로서 취급되어 활용가치가 적었다. 그러나 본 연구에서는 확률분포 모델을 적용하여 유용한 정보를 도출해 낼 수 있음을 확인하였다. 이러한 정보는 해당 작업장의 위생수준 평가, 동종 업체간의 위생수준 비교, 그리고 HACCP 적용시 적용전후의 개선상태 비교평가와 사후관리에도 활용 가능하다. 위해성평가(risk assessment)에서도 유용한 도구로 이용 가능한데, 예상되는 위해를 찾아내고 필요한 예방조치를 결정함에 있어서도 활용 가능한 것으로 판단된다[6].

FMC의 실제 활용에 있어서는 3가지의 기본조건을 전제로 수행되는 것으로 원칙으로 하고 있는데, 분석결과들이 독립적이어야 하고, 특별한 경향을 나타내지 말아야 하며, 일반세균수의 발생에 대해 동일한 환경적 또는 위생적 조건을 가지고 있어야 한다는 것이다[4]. 대부분의 식품산업체에서 수행하는 일반세균 검사자료는 특별한 위생관련 사고나 사건

Table 1. Distribution parameters of microbial counts in frozen dumplings with estimated and observed numbers of counts exceeding selected levels.

Source	Total no. of cases	Significant autocorrelation for lag(s)	Q-Q plot correlation coefficient	μ[1]	σ[2]	Nc[3]	P(N(n)= Nc) (%)[4]	Estimated No. in second half	Observed No. in second half
A	20	none	0.972	6.89	1.15	500	72.4	13~14	11
						1,000	49.6	9~10	11
						3,000	16.7	3	5
						5,000	7.9	1~2	2
						10,000	2.2	1	0
B	48	none	0.946	6.75	1.19	500	67.4	32~33	37
						1,000	44.7	21~22	36
						3,000	14.5	6~7	9
						5,000	6.8	3~4	3
						10,000	1.9	1	1

1) logarithms(LN) means(first half)
2) logarithms(LN) standard deviation(first half)
3) Nc : Threshold (CFU/g)
4) P(N(n)=Nc) : Probability of exceeding threshold

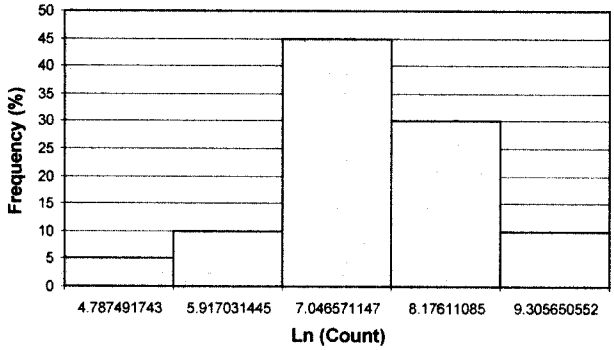

Fig. 7. The histograms of of the total microbial counts of frozen dumplings from source A.

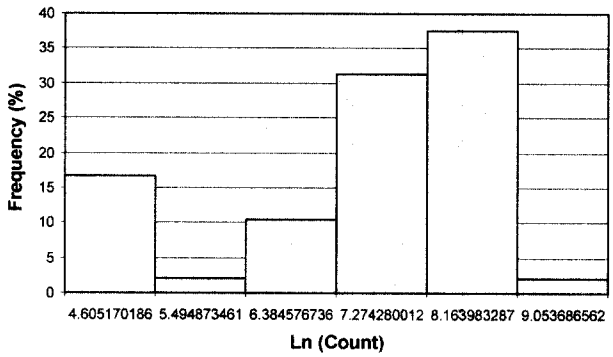

Fig. 8. The histograms of of the total microbial counts of frozen dumplings from source B.

을 경험하지 않았다면, 이러한 기본조건에 해당되는 자료인 것으로 판단된다.

충분한 수의 관찰 값(n)이 데이터베이스로 구축된다면 추정치의 정확성은 더 높아질 것이다. 본 연구에서는 관찰값이

Table 2. Classification of frozen dumplings sources according to the probability that total microbial count will exceed selected levels.

Rank	Source	n	Min.~Max. number (first half)	μ[1]	σ[2]	Nc[3]	P(N(n)= Nc) (%)[4]
1	B	48	100~8,550	6.75	1.19	10^5	0.296
2	A	20	120~11,000	6.89	1.15	10^5	0.298

1) logarithms(LN) means(first half)
2) logarithms(LN) standard deviation(first half)
3) Nc : Threshold (CFU/g)
4) P(N(n)=Nc) : Probability of exceeding

상대적으로 적다는 지적이 있겠으나, 과거 일반세균수의 분석에 대한 새로운 방법의 활용이라는 측면에서 충분한 가치가 있다고 하겠다. 또한 본 연구에서 이용한 확률분포모델인 Lognormal 분포이외에도 수집된 자료의 형태에 따라 다른 확률분포 모델의 적용도 가능할 것이다.

이러한 확률분포모델은 일반세균수외에 대장균군, 곰팡이, 효모 등의 불규칙적인 형태의 실험결과에도 적용 가능하다. 또한 목표기준을 초과할 확률을 계산하는데 이용할 수 있다. 실제 식품산업체에서도 일반 PC용 프로그램을 통해 쉽게 활용 가능하므로, 과거 이용 가치가 별로 없었던 일반세균수의 분포를 이용한다면 작업장내 안전성 관리에 매우 유익한 정보를 얻을 수 있을 것으로 판단된다.

감사의 말씀

이 논문은 2002년도 농림부 농림기술개발연구사업(202138031SB010)의 지원으로 이루어진 것이며, 이에 깊이 감사드립니다.

국문요약

식품산업체에서 검사한 자료의 일반세균수는 대부분 불규칙 분포를 나타낸다. 최근 이러한 불규칙적인 경향에 대하여 미생물수의 lognormal 분포 특성을 이용하는 확률분포 모델을 적용함으로써 좀 더 정확한 변화의 특성을 밝히고 있다. 확률분포 모델을 이용하면 일반세균수와 대장균군 등의 분석결과에 대해 과거의 단순한 경향분석을 벗어나, 유용한 분석이 가능한 것으로 나타나고 있다. 본 연구는 냉동만두류를 생성하는 2개 업체의 각각 1년 동안 실험된 일반세균 자료를 갖고, 확률분포 모델 활용의 타당성과 업체간의 위생수준 상호비교 그리고 예상되는 위해발생 예측에 적용 검토하였다. 그 결과 정량적인 미생물 자료에 대한 확률분포 모델의 적용은 식품산업체에서 일반세균수에 의한 위해발생 가능성을 예측하고 안전관리 수준을 결정하는데 유용한 것으로 판단되었다.

참고문헌

1. Horowitz, J., Normand, M.D. and Peleg, M.: On modeling the irregular fluctuations in microbial counts. *Critical Reviews Food Science Nutrition,* **39**(6), 503-517 (1999).
2. Corradini, M.G., Normand, M.D., Nussinovitch A., Horowitz, J. and Peleg, M.: Estimating the frequency of high microbial counts in industrial food products using various distribution functions. *J. Food Prot.*, **64**(5), 674-681 (2001).
3. Peleg, M., Nussinovitch, A. and Horowitz, J.: Interpretation and extraction of useful information from irregular fluctuating industrial microbial counts. *J. Food Science*, **65**, 740-747 (2000).
4. Peleg, M. and Horowitz, J.: On estimating the probability of aperiodic outbursts of microbial populations from their fluctuating counts. *Bulletin Mathematical. Biology*, **62**, 17-35 (2000).
5. Nussinovitch, A., Curasso, Y. and Peleg, M.: Analysis of the fluctuating microbial counts in commercial raw milk - A case study. *J. Food Protection*, **63**(9) 1240-1247 (2000).
6. Nussinovitch, A. and Peleg, M.: Analysis of the fluctuating patterns of microbial counts in frozen industrial food products. *Food Research International*, **33**, 53-62 (2000).
7. Peleg, M.: Estimating the Probability of High Microbial Counts Using Excel, http://www-unix.oit.umass.edu/~aew2000/microbecounts.html (2001).