

식중독 발생지수 개발

최국렬¹ · 김병수² · 배화수³ · 정우석⁴ · 조영준⁵

¹인제대학교 데이터정보학과; ²인제대학교 데이터정보학과; ³인제대학교 데이터정보학과;
⁴인제대학교 데이터정보학과; ⁵인제대학교 데이터정보학과

(2007년 12월 접수, 2008년 5월 채택)

요약

외식산업의 발달과 함께 학교와 기업 등의 단체 급식이 증가하고 있는 오늘날, 관련분야에서는 식중독 예방에 대해 많은 관심을 갖고 연구하고 있다. 우리나라에서는 온도에 따른 세균의 증식속도를 이용하여 식중독지수를 개발하여 식중독 위험에 대한 정보를 제공하고 있는데 그 정보가 식중독 발생상황과는 차이가 있음이 지적되고 있다. 본 연구에서는 최근 3년간(2004~2006년)의 일별 식중독 발생과 기상자료를 이용하여 일 최고기온, 습도, 월효과를 설명변수로 하고, 식중독 발생건수를 반응변수로 하는 로그선형모형(Log Linear Model)을 이용하여 식중독 발생의 위험을 예측할 수 있는 사고발생지수를 개발하였다. 개발된 지수와 기존지수를 비교한 결과 개발된 지수가 식중독 발생 상황을 반영함에 있어서 더 나은 설명력이 있음이 나타났다.

주요용어: 식중독, 로그선형모형, 식중독발생지수.

1. 서론

외식산업이 발달하고 학교와 기업 등의 단체 급식이 증가함에 따라 (김진수, 2006) 우리나라에서도 발생하는 식중독 발생건당 규모가 커지고 발생건수에서도 증감을 반복하며 일정수준으로 지속적으로 발생하고 있다 (그림 1.1). 식중독이 발생하게 되면 증상의 차이는 있지만 일반적으로 많은 사람이 고통을 받고 정상적인 활동에 제약을 받게 되므로 미국을 비롯한 선진국에서는 식중독을 예방하기 위해 교육, 홍보 그리고 연구를 통해 대책을 마련하고 지속적인 관리를 통해 식중독을 감소시키기 위해 힘쓰고 있으며 (Krupnick 등, 2005) 이러한 경향은 선진국형의 식생활 문화로 바뀌어 가고 있는 우리나라에서도 예외는 아니다 (신광순 등, 2004).

우리나라에서는 식품의약품안전청이 식중독 예방에 대한 주관업무를 맡아서 식중독 예방을 위한 연구를 하고 대국민 홍보사이트를 이용하여 관련 국내외 소식, 식중독에 관한 자세한 이해 및 예방법등에 대해 홍보와 함께 관련 종사자 교육을 실시함으로써 식중독 발생에 관한 전반적인 관리를 하고 있다 (권기성,

본 연구는 2007 식품의약품안전청 용역연구개발사업(과제번호:07032기타사743)의 지원에 의해 이루어졌음.

¹(621-749) 경남 김해시 어방동 607, 인제대학교 데이터정보학과, 통계정보연구소 교수.

E-mail: choi@stat.inje.ac.kr

²(621-749) 경남 김해시 어방동 607, 인제대학교 데이터정보학과, 통계정보연구소 조교수.

E-mail: kbs@stat.inje.ac.kr

³(621-749) 교신저자. 경남 김해시 어방동 607, 인제대학교 데이터정보학과, 통계정보연구소 교수.

E-mail: wbae@stat.inje.ac.kr

⁴(621-749) 경남 김해시 어방동 607, 인제대학교 데이터정보학과, 석사과정. E-mail: 99224029@hanmail.net

⁵(621-749) 경남 김해시 어방동 607, 인제대학교 데이터정보학과, 석사과정. E-mail: choyj1004@nate.com

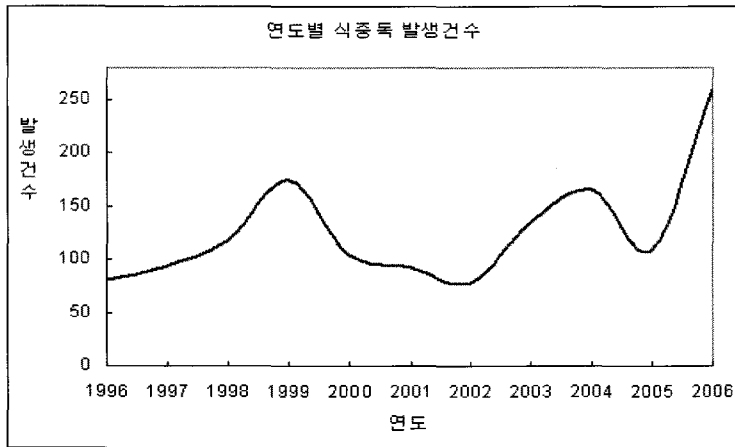


그림 1.1. 우리나라 연도별 식중독 발생건수

2006; 최순곤, 2007). 특히 기상청과 연계하여 개발된 식중독지수는 생활지수 중 하나로 식중독 위험에 대한 정보로 제공되고 있다. 일 최고기온을 기준으로 계산하도록 개발되어 있는 이 지수는 온도에 대한 증가함수로 기온이 높을수록 지수가 높아져서 위험률에 대한 예보의 수치도 높아지게 되지만 실제 발생 상황과는 차이가 난다는 것이 문제가 되고 있다.

본 연구에서는 기존지수를 보완하고 식중독 발생상황을 더 잘 설명할 수 있는 지수를 개발하기 위해 식중독 발생건수를 반응변수로 하고 설명변수로는 일 최고기온과 상대습도 그리고 월효과를 사용하여 통계적 모형을 적합시키고 이로부터 식중독 사고발생지수를 개발하게 되었다.

제 2장에서는 기존 식중독지수에 대하여 정리를 하고 제 3장에서는 새로운 지수의 개발 과정에 대해 다루고, 제 4장에서는 개발된 지수와 기존지수를 비교하고 결론과 제언을 제 5장에 다루었다.

2. 기존지수

기존지수는 주어진 온도에 대한 식중독위험에 대한 예보지수로 식품의약품안전청과 기상청이 함께 개발하여 제공하고 있다. 이는 세균증식의 최적조건에서의 식중독유발시간과 해당온도에서의 식중독 유발시간에 대한 비율로 나타내며 0에서 100사이의 값을 가진다.

기존지수 개발에서의 식중독 유발시간 = 증식시간 \times 10으로 세균의 10회 증식까지 걸리는 시간을 나타내고 증식시간 = $\ln(2) \div$ 성장속도 = $0.693 \div$ 성장속도 = 1회 분열 시간을 나타내는데 이는 Ingraham (1958)에 의해 연구된 E. coli에서의 온도와 성장속도 간의 관계를 이용하여 계산된다. 균이 증식할 수 있는 최적조건은 온도 40°C, pH=6.5~7.0, 수분활성도=1~0.99를 의미하며 이 조건에서 초기 균수 1,000마리인 식품이 식중독을 발생시킬 수 있는 1,000,000개의 균수로 증식되는데 걸리는 시간은 3.5시간이 소요되는데 기존지수 100이라 함은 식중독 유발시간이 3.5시간인 발생 최적의 상태를 의미한다. 기존지수는 균 증식의 조건 중, 온도에 대한 조건만을 일 최고기온으로 하고 온도 외의 pH와 수분활성도는 최적인 상태라고 가정하여 개발되었으며 정수인 값의 온도에 대해서만 지수가 제공되고 있다 (표 2.1). 본 논문에서는 기존지수의 값을 개발된 지수와 비교함에 있어서 연속된 온도에서의 지수의 값이 필요하므로 정수에서 계산된 기존지수의 값을 이용하여 Xu 등 (2001)이 제안한 커널 추정법(kernel estimation method)을 사용하여 기존 지수의 값을 추정하였다. 기존지수의 값은

표 2.1. 최고기온에 따른 기준지수

최고기온 (°C)	성장속도 (hr ⁻¹)	증식시간 (hr)	식중독 유발시간	식중독 지수
40	1.98	0.35	3.5	100
39	1.95	0.36	3.6	98
38	1.92	0.36	3.6	97
37	1.85	0.37	3.7	93
36	1.80	0.39	3.9	91
35	1.70	0.41	4.1	86
34	1.60	0.43	4.3	81
33	1.45	0.48	4.8	73
32	1.35	0.51	5.1	68
31	1.20	0.58	5.8	61
30	1.05	0.66	6.6	53
29	1.00	0.69	6.9	51
28	0.90	0.77	7.7	45
27	0.80	0.87	8.7	40
26	0.70	0.99	9.9	35
25	0.60	1.16	11.6	30
24	0.55	1.26	12.6	28
23	0.475	1.46	14.6	24
22	0.40	1.73	17.3	20
21	0.35	1.98	19.8	18
20	0.25	2.77	27.7	13
19	0.20	3.47	34.7	10
18	0.18	3.85	38.5	9
17	0.15	4.62	46.2	8
16	0.12	5.78	57.8	6
15	0.10	6.93	69.3	5

$\hat{y} = \hat{f}(x) = \sum_{i=1}^n \alpha_i \exp(-(x - x_i)^2 / 2\sigma^2) + b$ 로 추정되는데 여기서 \hat{y} 은 지수계산을 원하는 온도 x 에서의 기준지수의 추정값, α_i 는 주어진 온도 x_i 에서 MSE(Mean Squared Error)를 최소화시키는 가중치, x_i 는 자료에 주어진 온도 그리고 $n = 26$ 은 자료에 주어진 관측치의 수를 나타내고 b 는 편의(bias), σ 는 평활량(bandwidth)을 나타내는데 이 자료에서는 b 는 64.1132, σ 는 5를 사용하였다 (표 2.2).

그림 2.1은 2004년부터 2006년까지 3년간의 서울/경기/인천지역의 세균성 식중독에 대한 월별 평균 상황을 나타내고 있다. 2004년부터 2006년까지 3년간의 서울/경기/인천지역의 세균성 식중독에 대한 자료 분석 결과, 기온이 낮은 달의 발생률은 낮고 기온이 높은 달에는 높게 나타나고 있음을 알 수 있다. 일최고기온이 가장 높은 8월의 경우엔 기준지수가 최고로 나타나고 있으나, 발생률은 6월에 제일 높게 나타났고, 7월에도 높은 기준지수로 보아 발생률이 높을 것으로 예상되나 5월의 발생률과 비슷하게 나타났다. 또한 1~3월에는 최고기온의 변화 및 발생률의 변화가 있음에도 불구하고 기준지수는 5근방으로 거의 일정하게 표시되고 있어 기준 지수는 보완이 필요하다는 것을 알 수 있다.

표 2.2. 추정식의 변수

x	a	x	a
15	-84.7833	28	-4.6653
16	-84.7833	29	-4.6653
17	-84.7833	30	-4.6653
18	-84.7833	31	-4.6653
19	-84.7833	32	-4.6653
20	-84.7833	33	-4.6653
21	-84.7833	34	-4.6653
22	-84.7833	35	-4.6653
23	-84.7833	36	-4.6653
24	-84.7833	37	-4.6653
25	-84.7833	38	-4.6653
26	-84.7833	39	-4.6653
27	-84.7833	40	-4.6653

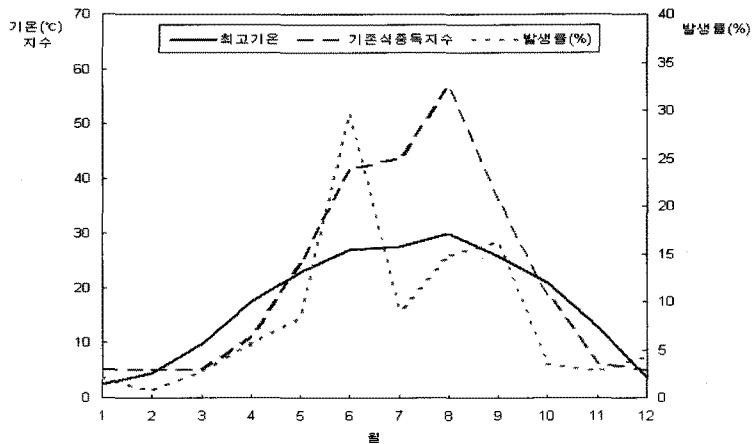


그림 2.1. 월별 최고기온과 식중독지수, 발생률의 시계열도(서울/경기/인천)

3. 모형개발

모형개발을 위해 사용한 자료는 3년간(2004년~2006년)의 일별 식중독 발생건수(출처: 식품의약품안전청 제공)와 기상자료(출처: 기상청 제공)를 결합하여 이루어진 1,096개의 관측치로 구성된다. 식중독에 관련된 원인물질로 여러 종류가 있으나 주된 원인이 세균에 의한 것이므로 세균에 의한 식중독의 발생에만 관심을 두기로 한다. 또한 우리나라의 절반가량의 식중독이 서울/경기/인천 지역에서 발생함이 최근 3년의 자료에 나타났으므로, 서울/경기/인천지역의 세균에 의한 일별 식중독 발생건수를 반응변수로 채택하고, 설명변수로 사용되는 기상자료는 서울관측소의 기상관측 자료를 서울/경기/인천지역의 대표값으로 하여 모형개발에 사용하기로 한다.

표 3.1. 모형에 대한 적합도

Criterion	Df	Value	Value/DF
Deviance	1082	587.0248	0.5425
Scaled Deviance	1082	587.0248	0.5425
Pearson Chi-Square	1082	1355.9769	1.2532
Scaled Pearson χ^2	1082	1355.9769	1.2532
Log Likelihood		-376.3979	

표 3.2. 모수 추정 및 검정

Parameter	Df	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	-3.248	0.7575	18.38	<.0001
temp _{max}	1	0.028	0.0267	1.12	0.2904
hum	1	0.004	0.0072	0.38	0.5378
m1	1	-0.511	0.6771	0.57	0.4508
m2	1	-1.568	0.9870	2.52	0.1122
m3	1	-0.418	0.5233	0.64	0.4247
m4	1	0.091	0.3539	0.07	0.7975
m5	1	0.287	0.3289	0.76	0.3837
m6	1	1.421	0.3378	17.70	<.0001
m7	1	0.148	0.4436	0.11	0.7394
m8	1	0.592	0.4414	1.80	0.1798
m9	1	0.859	0.3344	6.60	0.0102
m10	1	-0.540	0.4410	1.50	0.2209
m11	1	-0.504	0.4885	1.06	0.3022

3.1. 모형 추정

일별 식중독 발생건수를 반응변수 Y 로 두면 Y 는 계수형 변수로 포아송 분포 가정에 무리가 없어 포아송 확률분포를 이용하는 로그선형모형을 이용하기로 한다 (Agresti, 2007; Lindsey, 1995). 설명변수로는 세균의 증식 조건 중, 온도와 수분활성도와 관계되는 일 최고기온과 상대습도를 선택하고 이 조건에서 예측할 수 있는 상황과 실제 발생상황에 대한 차이를 보완하기 위해 월 효과를 추가하기로 한다. 겨울과 봄 동안에는 식중독에 안일하게 대처하다가 6월경에는 장마나 온난화로 인한 기온의 상승 등으로 무방비상태에서 식중독이 발생하게 되고 이로부터는 매우 조심을 하기 때문에 7, 8월에는 균의 증식이 높은 조건에도 불구하고 발생이 낮아진다고 볼 수 있으므로 이를 월효과로 모형에 나타낼 수 있도록 한다.

반응변수인 일별 식중독 발생건수 Y 가 평균이 μ 인 포아송분포를 따른다고 가정하고 연결함수로 $\log(\mu)$ 를 사용하는 로그선형모형은 식 (3.1)과 같이 나타낼 수 있다.

$$\log(\mu) = \beta_0 + \beta_1 \text{temp}_{max} + \beta_2 \text{hum} + \sum_{i=1}^{12} \beta_{3i} m_i, \tag{3.1}$$

여기서 temp_{max} 와 hum 은 각각 일 최고기온과 일 상대습도를 나타내는 설명변수이고

$$m_i = \begin{cases} 1, & i\text{월} \\ 0, & \text{otherwise} \end{cases}, \quad i = 1, 2, \dots, 12 \text{를 나타내며 } m_i \text{는 가변수로 사용된다.}$$

표 3.3. \hat{p} 의 분포

변수	N	평균값	표준편차	최소값	최대값
\hat{p}	1096	0.1150	0.1008	0.0075	0.4247

추정된 모형으로부터 일 평균 식중독 발생건수는 $\hat{\mu} = \exp(\hat{\beta}_0 + \hat{\beta}_1 \text{temp}_{max} + \hat{\beta}_2 \text{hum} + \sum_{i=1}^{12} \hat{\beta}_{ii} m_i)$ 를 이용하여 추정할 수 있는데, 자료를 이용하여 모형을 추정한 결과, $\hat{\mu} = \exp(A)$, 여기서 A 는 다음과 같이 구해진다.

$$\begin{aligned} A = & 3.248 + 0.028\text{temp}_{max} + 0.004\text{hum} - 0.511m_1 - 1.568m_2 \\ & - 0.418m_3 + 0.091m_4 + 0.287m_5 + 1.421m_6 + 0.148m_7 \\ & + 0.592m_8 + 0.859m_9 - 0.540m_{10} - 0.504m_{11} + 0.143m_{12}. \end{aligned} \quad (3.2)$$

모형적합도(goodness of fit)에 대한 결과는 표 3.1에 그리고, 모수 추정 및 검정 결과는 표 3.2에 나타난 바와 같다.

표 3.2의 월효과($m_1 \sim m_{11}$)는 12월의 월효과인 m_{12} 에 대한 상대적인 크기에 대한 검정결과를 나타내는데 6월과 9월의 효과가 유의하게 나타나는데 이는 적합된 모형이 그림 1.1에 나타난 발생상황을 잘 반영하고 있음을 보여준다.

3.2. 지수 개발

모형으로부터 특정한 하루에 식중독이 1건 이상 발생할 확률을 추정하면

$$\begin{aligned} \hat{p} &= P(Y \geq 1 | \hat{\mu}) \\ &= 1 - \Pr(Y = 0 | \hat{\mu}) \\ &= 1 - \exp(-\hat{\mu}) \\ &= 1 - \exp(-\exp(A)) \end{aligned}$$

과 같이 나타내어지고, A 는 식 (3.2)에 나타난 바와 같다.

자료로부터 구한 최근 3년의 일별 식중독발생확률의 분포를 보면 표 3.3과 그림 3.1과 같은데 평균이 0.11이고 발생확률이 0.1이하가 거의 60%에 해당하며 왼쪽으로 많이 치우친 분포를 나타내므로 식중독은 흔히 발생하지 않는다는 사실을 알게 된다.

위에서 얻은 \hat{p} 을 이용하여 100을 곱하여 지수를 개발하게 되면 너무 작은 값으로 나타나므로 위험도를 반영하기엔 충분하지 못하게 된다. 즉, $\hat{p} \times 100$ 을 이용한 경우의 40은 매우 위험하지만 100을 기준으로 하면 상당히 안전해 보인다. 따라서 기준지수와 같은 범위로 근접하게 맞추고 조정시키기 위해 \hat{p} 의 값을 표준화한 다음 척도변환을 하여 전체 범위 100에서 평균이 50근방인 46 그리고 표준편차는 15를 갖는 분포를 만들어 이를 사고발생지수(index)라 한다.

$$\text{index} = \hat{p} \times 148.81 + 28.89. \quad (3.3)$$

사고발생지수는 식중독발생 현황자료를 이용하여 얻어진 당일 최고기온, 습도의 기상조건과 월효과를 반영하여 개발된 식중독 발생에 대한 예측지수로 정의될 수 있다. 서울/경기/인천지역에 대해 2004년부터 2006년까지 자료를 이용하여 사고발생지수를 구하고 그 분포를 살펴보면, 표 3.4와 그림 3.2에 나

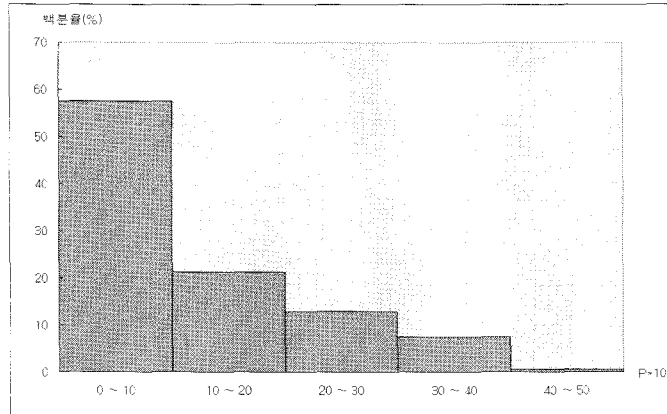


그림 3.1. p 의 분포

표 3.4. 사고발생지수의 통계량(서울/경기/인천, 2004년~2006년)

변수	N	평균값	표준편차	최소값	최대값
p	1096	46.00	15.00	30.00	92.08

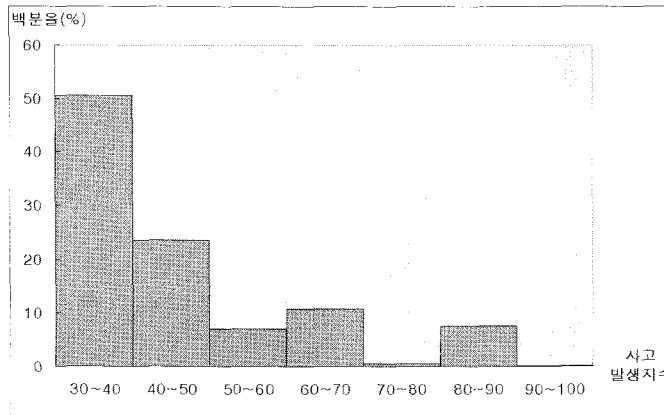


그림 3.2. 사고발생지수의 분포(서울/경기/인천, 2004년~2006년)

타나 있는 것처럼 평균이 46.00, 최소값이 30.00, 최대값이 92.08로 나오고 있어 지수가 나타내는 값과 위험도가 어느 정도 병행하게 지수의 범위가 맞춰짐을 볼 수 있다. 그림 3.2에 나타난 지수의 분포를 보면 3년 동안 지수 30이하는 없고, 지수가 30~40 사이가 50%로 가장 높으며 50이하가 75%정도를 차지하고 90~100 사이인 일수가 0.08%이고 80~90 사이가 7.57%를 차지하고 있다.

계산된 사고발생지수는 나타난 값 그대로 식중독 발생에 대한 위험을 예측하게 하는 지표로 활용될 수 있는데 이 지수가 개발된 과정을 역으로 환산하면 주어진 지수에서 예측하게 되는 일 평균 식중독 발생 건수도 계산할 수 있어 식중독 발생에 대한 정보로 병행할 수 있다. 즉, $지수 = \hat{p} \times 148.81 + 28.89$ 를 이 용해서 구해졌으므로 이 식을 이용하면 $\hat{p} = (지수 - 28.89) / 148.81$ 로 구해지는데 $\hat{p} = \Pr(Y \geq 1 | \hat{\mu} = 1 - \exp(-\hat{\mu}))$ 이므로 $\hat{\mu} = -\log(1 - \hat{p})$ 으로 구해진다. 따라서 각 지수에 대해 $\hat{\mu}$ 이 계산되므로 이를 통

표 3.5. 2005년 인구조사 결과

	전국	서울/경기/인천	기타지역
인구수	47,041,434	22,621,232	24,420,202

표 3.6. 지수의 예보단계

단계	지수		\hat{p}		$\hat{\mu}$ (서울기준)		$\hat{\mu}$ (전국기준)	
	하한	상한	하한	상한	하한	상한	하한	상한
4	81	100	0.3502	0.4779	0.4312	0.6498	0.8969	1.3516
3	70	81	0.2763	0.3502	0.3233	0.4312	0.6725	0.8969
2	58	70	0.1956	0.2763	0.2177	0.3233	0.4528	0.6725
1	0	58	0.0000	0.1956	0.0000	0.2177	0.0000	0.4528

표 4.1. 발생건수와의 상관관계

상관계수	기존지수	개발된 지수
발생건수(전체)	0.19 (<.0001)	0.26 (<.0001)
발생건수(5월~9월)	0.04 (<.3753)	0.18 (<.0001)

해 주어진 지수에서 예측할 수 있는 일평균 발생건수를 정보로 이용할 수 있다.

또한 얻어진 모형은 서울/경기/인천 지역의 자료를 이용하였으므로 전국 지역에 대한 일 평균발생건수에 대한 예측을 하기 위해 인구에 대한 정보를 이용하기로 한다. 표 3.5에 의해서 전국과 서울/경기/인천 지역의 인구비 즉, $47,041,434/22,621,232 = 2.080$ 를 이용하면 $\hat{\mu}$ (전국)은 $\hat{\mu}$ (서울/경기/인천 기준) $\times 2.080$ 을 이용하면 구해진다. 전국의 세균에 의한 식중독 발생건수에 대한 예측값과 기존지수의 값을 참고로 하여 표 3.6에 나타낸것과 같이 예보단계를 나눌 수 있다.

4. 지수검토

개발된 지수를 검토하기 위해서 발생건수와의 상관관계를 비교해 보면 (표 4.1), 개발된 지수와는 상관계수가 0.26으로 나타났고, 기존지수와는 0.19로 개발된 지수보다 낮게 나타났다. 특히 식중독 발생이 많은 5월에서 9월 사이의 기간에서 발생건수와의 상관관계를 보면 개발된 지수와는 0.18이고, 기존지수가 0.04로 기존지수는 거의 발생에 대한 예측을 하지 못하는 반면, 개발된 지수는 발생상황과 연계성이 더 높음을 알 수 있다.

그림 4.1의 기존지수와 사고발생지수, 발생건수를 3년 동안 월평균으로 본 시계열도에서도 6월에서의 발생을 기존지수가 예측하지 못한 반면, 사고발생지수는 발생상황에 부합하고 있음을 알 수 있다.

5. 결론 및 제언

기존지수는 온도에 따른 세균의 증식시간만을 반영하여 개발되어 있어서 실제 발생 상황을 예측하는데 있어서 예측력이 부족하다는 점이 지적되어 왔다. 본 연구에서는 이를 보완하기 위하여 지난 3년간의 서울/경기/인천 지역의 세균성 식중독 발생자료를 기반으로 일 최고기온과 함께 월 효과와 상대습도를 고려하고 식중독 발생건수를 반응변수로 하여 포아송 확률분포를 이용한 로그선형모형을 적합시켰다.

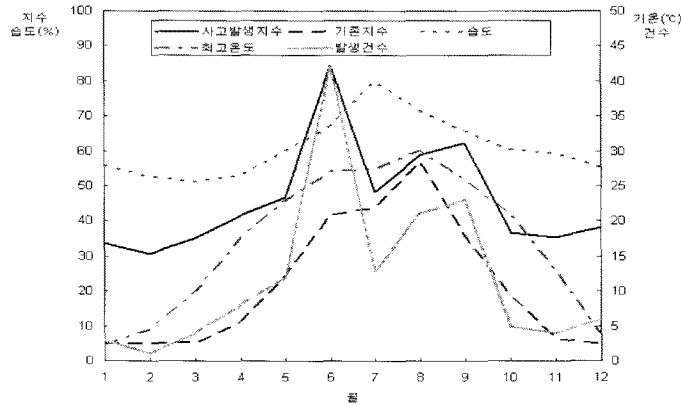


그림 4.1. 사고발생지수, 기본지수의 시계열도(서울/경기/인천, 2004년~2006년의 월평균)

개발된 모형으로부터 식중독이 1건 이상 발생할 확률을 근거로 사고발생지수를 개발하여, 기존지수의 범위와 근접하도록 변환하여 전국의 식중독 일평균발생건수를 예측하였고, 예측 값에 의해 네 단계로 예보단계를 나누었다. 기존지수와 사고발생지수에 대한 발생건수와의 상관분석에서 개발된 지수는 기존지수보다 두 배 이상의 상관관계를 보이며 발생상황을 잘 설명하였으며 지난 3년간의(2004년~2006년) 자료를 통해서 나타난 월별 발생상황에서도 기존지수가 설명하지 못한 6월과 9월의 발생상황을 잘 설명해주고 있음을 보였다. 식중독 발생은 흔히 일어나는 경우(평균 10% 전후)가 아니므로 발생에 대한 자료가 부족하고 원인물질이 불명(30% 가량)인 자료가 많아서 모형 구축에 어려움이 있었으나 본 연구에서는 주된 원인물질인 세균으로 국한시켰다. 기상자료에 대해서도 전국을 대상으로 하면 자료의 양도 많아지고 모형 구축에 도움이 될 수 있으나 기상자료의 대표값 선정에 대한 방법론에서 기상청에서도 뚜렷하게 제시하지 못해서 서울/경기/인천 지역의 자료로 서울 관측소의 값을 사용하였다. 기상변수에 대한 전국 대푯값을 선정하는 합리적인 방법을 개발하여 전국의 자료를 사용하면 전국의 발생패턴을 보다 더 잘 설명할 수 있는 모형과 지수개발이 가능해 질 것이라고 본다.

참고문헌

권기성 (2006). 식중독의 발생동향 및 관리체계, <보건복지포럼>, 118, 17-25.
 김진수 (2006). 단체급식과 식품안전정책방향, <식품산업과 영양>, 11, 35-37.
 신광순, 광동경, 문혜경, 이지현, 이정화, 김윤주 (2004). 선진국의 식중독 관리 시스템 조사, <식품의약품안전청 연구보고서>, 8, 2091-2092.
 최순곤 (2007). 식중독 예방관리 대책 추진 방향, <광주지방식품의약품안전청 식중독예방 Workshop>.
 Agresti, A. (2007). *An Introduction to Categorical Data Analysis*, John Wiley & Sons. Inc. Hoboken. New Jersey.
 Ingraham, J. L. (1958). Growth of psychrophilic bacteria, *Journal of Bacteriology*, 76, 75-80.
 Krupnick, A., Batz, M. and Hoffmann, S. (2005). Prioritizing opportunities to reduce the risk of foodborne illness: A conceptual framework, *Discussion Paper Series of Food Safety Research Consortium*.
 Lindsey, J. K. (1995). *Modelling Frequency and Count Data*, Clarendon Press, Oxford.
 Xu, J., Zhang, X. and Li, Y. (2001). Kernel MSE algorithm : A unified framework for KFD, LS-SVM and KRR, In *Proceedings of the International Joint Conference on Neural Networks (IJCNN-2001)*, 2, 1486-1491.

Developing the Index of Foodborne Disease Occurrence

Kookyeol Choi¹ · Byungsoo Kim² · Whasoo Bae³ · Wooseok Jung⁴ · Youngjoon Cho⁵

¹Dept. of Data Science, Inje University; ²Dept. of Data Science, Inje University;

³Dept. of Data Science, Inje University; ⁴Dept. of Data Science, Inje University;

⁵Dept. of Data Science, Inje University

(Received December 2007; accepted May 2008)

Abstract

As the Eating Out Businesses are making rapid progress and most of the schools and the firms serve the meals, the foodborne disease has occurred increasingly and lots of researches and the policies are studied to prevent it. In Korea, the foodborne disease index for prevention is developed by using bacterial growth rate on the temperature to give the information about the danger level of the foodborne disease, but the gap between real status of the occurrences and the predicted danger level has been pointed out. This study aims at developing the index of the foodborne occurrence based on the log linear model using the data of the foodborne disease occurrence and the meteorological data for the last three years(2004 ~ 2006). Comparison between the new index and the existing index showed that the new index is better in explaining the foodborne disease occurrence.

Keywords: Foodborne disease, log linear model, index of the foodborne disease occurrence.

This work was supported by Korea Food & Drug Administration R&D Project in 2007.

¹Professor, Dept. of Data Science/Institute of Statistical Information, Inje University, Gimhae, Kyungnam 621-749, Korea. E-mail: choi@stat.inje.ac.kr

²Assistant Professor, Dept. of Data Science/Institute of Statistical Information, Inje University, Gimhae, Kyungnam 621-749, Korea. E-mail: kbs@stat.inje.ac.kr

³Corresponding Author: Professor, Dept. of Data Science/Institute of Statistical Information, Inje University, Gimhae, Kyungnam 621-749, Korea. E-mail: wbae@stat.inje.ac.kr

⁴Graduate Student, Dept. of Data Science, Inje University, Gimhae, Kyungnam 621-749, Korea. E-mail: 99224029@hanmail.net

⁵Graduate Student, Dept. of Data Science, Inje University, Gimhae, Kyungnam 621-749, Korea. E-mail: choyj1004@nate.com