

빅 데이터 분석을 통한 기상 요인과 식중독 발병의 연관성

식품영양학과 19
2019101918 조민서

목차

1. 서론

2. 본론

2.1 연구 방법

2.1.1 기상 자료

2.1.2 식중독 발생 자료

2.2 분석 방법

2.2.1 랜덤 포레스트

2.2.2 로지스틱 회귀

2.3 변수 선택

3. 연구 결과

4. 고찰 및 결론

요 약

본 연구는 서울시의 2018년 1월 1일부터 2018년 12월 31일까지의 기상 변이에 관한 빅 데이터 자료와 식중독 발생사고 빅 데이터 자료를 융합하여 식중독 발병률 변이와 기상요인의 관계에 대한 분석을 시도하여 국민 건강 예방에 도움을 주고자 한다. 이 연구는 기존의 “기상 요인과 식중독 발병의 연관성에 대한 연구”를 기반으로 진행되었으며 식중독 원인물질에 식품 위생학 전공 내용을 추가하여 연구를 진행하였다. 이번 연구에서는 분석 도구 R을 이용하여 랜덤 포레스트로 변수 축약을 하고 로지스틱 회귀로 데이터를 분석하였다. 로지스틱 회귀 분석 결과, 월합 강수량, 평균풍속, 평균최저기온이 높을수록 감염형 세균성 원인균에 의한 식중독 발병률이 높게 나타나며 최저 해면기압, 최대 풍속이 낮을수록 감염형 세균성 원인균에 의한 식중독 발병률이 높게 나타났다. 최저 상대습도, 일 최다 강수량, 일사량, 평균최저기온이 높을수록 독소형 세균성 원인균에 의한 식중독 발병률이 높게 나타나며, 최저 해면기압, 평균 풍속, 월 강수량 합이 낮을수록 독소형 세균성 원인균에 의한 식중독 발병률이 낮게 나타났다. 월합강수량, 일사량, 최저기온이 높을수록 독소 감염형 세균성 원인균에 의한 식중독 발병률이 높게 나타나며, 최저 해면기압, 최저 상대습도, 최대 풍속이 낮을수록 독소 감염형 세균성 원인균에 의한 식중독 발병률이 낮게 나타났다. 바이러스에 의한 식중독 발병률은 최저 해면기압, 평균풍속이 높을수록 높게 나타나며, 평균, 최저 상대습도, 일 최다 강수량, 일사량, 최저기온이 낮을수록 바이러스성 원인균에 의한 식중독 발병률이 높게 나타났다. 크게 두 가지 원인균인 세균과 바이러스에 의한 식중독 발병률이 같은 기상요인에 의해 영향을 받더라도 세균성 원인균에 의한 식중독 발병률은 정(+)의 영향을 바이러스성 원인균에 의한 식중독 발병률은 부(-)의 영향을 받는 것으로 나타났다. 또한 세균성 원인균들 중 독소형, 독소 감염형 세균은 감염형 세균과 다르게 일사량에도 큰 영향을 받는 것을 확인할 수 있었다.

1. 서론

해마다 여름이 되면 음식 등에 세균 증식이 활발해지면서 식중독에 걸리는 경우가 많아진다. 여름뿐만 아니라 겨울에도 노로바이러스로 인한 식중독 사고가 빈번하게 일어나는 것을 확인할 수 있다. 우리는 사계절 내내 식중독 균과 바이러스에 노출될 위험성을 내재하고 있기 때문에 식품섭취에 유의해야 한다. 식중독과 같은 물이나 음식에 의한 매개질환들은 기후 변수와 밀접한 관계가 있는 것으로 알려져 있다.[6] 식품의약품안전처의 식중독 발생 현황 자료를 통해, 우리나라의 경우 최근 10년간 끊임없이 식중독이 발생하고 있는 것을 확인할 수 있다. 이 자료를 통해 세균성 식중독은 주로 여름철, 바이러스성 식중독은 주로 겨울철에 발생하는 양상을 확인할 수 있다. 이 자료를 통해서도 식중독은 기후와 연관성이 있다는 것을 짐작해볼 수 있다.

이 연구는 기존의 “기상요인과 식중독 발병의 연관성에 대한 연구”자료를 토대로 진행되었다. 기존의 연구 자료는 식중독을 세균성 식중독과 바이러스 식중독으로 구분하였다. 필자는 식품영양학과 학생으로서 기존의 연구보다 식중독 원인물질을 감염형 세균, 독소형 세균, 독

소 감염형 세균, 바이러스에 의한 식중독으로 좀 더 세분화하였다. 식중독 발생 사고 시 대부분의 사람들은 식중독 발생 원인물질을 세균, 바이러스 원인물질 정도로만 알고 있다. 하지만 세균성 원인물질에도 각 세균마다 성장 조건, 증상 등이 다르다. 감염형 세균은 적은 양으로도 증식이 가능하고 잠복기가 길다는 특징이 있다. 반면 독소형 세균은 잠복기가 짧고 장독소를 만들어내어 복통, 설사, 구토뿐만 아니라 마비, 탈수, 심각한 경우 사망까지 이르게 할 수 있다. 독소 감염형 세균은 잠복기가 8시간~24시간으로 감염형 세균과 독소형 세균의 중간 정도이며, 포자를 형성하여 낮은 pH, 높은 열, 건조한 환경에서도 잘 성장한다는 특징을 가지고 있다. 그러므로 독소형, 독소 감염형 세균은 감염형 세균보다 치명적인 독소를 생성하기 때문에 더 위험하다고 할 수 있다. 독소형, 독소 감염형 세균은 내열성이 강하고 포자를 형성하여 환경에 저항하는 힘이 강하기 때문에 세균이 번식하기 좋은 6월~9월에는 이러한 세균을 예방하는 방법을 알고 있는 것이 중요하다고 생각된다.[7]

식중독을 일으키는 바이러스 원인물질은 세균성 원인물질과 다르게 11월~3월까지 겨울에 증식되는 경향을 확인할 수 있다. 바이러스는 세균과 다르게 2차 감염이 가능하기 때문에 집단적인 발병 양상이 나타난다. 또한 세균성 원인물질의 일부 균은 백신이 개발되어 있지만 바이러스는 일반적 치료법이나 백신이 없기 때문에 바이러스 식중독에 대한 경각심이 필요하다.[7]

식중독 원인물질들의 큰 차이점 중 하나는 각 물질마다 식중독 사고가 일어나는 계절에 있다. 세균성 원인물질에 의한 식중독 사고는 주로 여름철, 바이러스성 원인물질에 의한 식중독 사고는 주로 겨울철에 일어나는 양상을 확인할 수 있다. 또한 독소 감염형인 *Clostridium perfringens*는 봄철에 자주 발생되고 있어 세균성 식중독 원인물질 안에서도 감염형, 독소형 등의 세균을 계절별로 유의해야 할 필요가 있다.[7]

병원균은 특정 온도에서 생존 및 증식하며, 기온 증가는 병원균의 외인자 잠복기와 번식에 영향을 미친다. 특히, 추운 날씨와 낮은 습도는 바이러스를 증식 시키는 것으로 알려져 있다.

기온과 강수량은 식품 매개질환, 연안 해안 수질 등과 밀접한 관계가 있고, 비브리오 등과 같은 식품 매개질환에 미치는 기후변화의 영향은 매우 크다는 국내 연구결과들이 보고되고 있다.[3]

유럽 몇몇 국가들의 기후 변화와 감염형 세균인 살모넬라에 의한 식중독 발생의 연관성이 존재함을 확인하였으며[4], 특히, 여름철 수온 상승과 감염형 세균인 비브리오균에 의한 감염은 강한 연관성이 존재하는 것으로 확인되었으며 비브리오균에 의한 감염은 5월과 8월 사이에 78%의 발병률을 가지는 것으로 보고되었다.[5]

여러 연구결과에서와같이 질병에 영향을 미치는 원인균들은 기상 변이와 상관관계가 높다는 것을 알 수 있다.

본 연구에서는 기온, 상대습도, 기압, 풍속 등을 포함한 기상요인들과 식중독 발생의 상관성을 확인한다. 이 연구를 기반으로 한 “기상요인과 식중독 발병의 연관성에 대한 연구”에서는 식중독을 서로 다른 성격을 가지는 세균성 원인균에 의한 식중독과 바이러스성 원인균에 의한 식중독으로 구분하였다. 이 연구는 좀 더 심화적인 내용을 추가하여 세균성 원인균을 감염형 세균성 원인균, 독소형 세균성 원인균, 독소 감염형 세균성 원인균으로 구분하여 기상에 의한 식중독 발병률 변이를 확인하고자 한다.

2. 연구방법

2.1 연구자료

2.1.1 기상 자료

이 연구는 기상청의 서울시의 2018년 1월 1일부터 2018년 12월 31일까지의 월 단위 지상 관측 빅 데이터 자료를 사용하였으며, 종관기상관측(ASOS) 자료와 방재기상관측(AWS) 자료를 이용하였다. ASOS 자료는 종관 규모의 날씨를 파악하기 위하여 정해진 시각에 모든 관측소에서 같은 시각에 실시하는 지상관측을 말한다. 종관 규모는 일기도에 표현되어 있는 보통의 고기압이나 저기압의 공간적 크기 및 수명을 말하며, 주로 매일의 날씨 현상을 뜻한다. AWS는 지진, 태풍, 홍수, 가뭄 등 기상현상에 따른 자연재해를 막기 위해 실시하는 지상관측을 말한다. 이 자료는 관측 공백 해소 및 국지적인 기상 현상을 파악하기 위하여 서울시 28개 지점에 자동기상관측장비(AWS)를 설치하여 자동으로 관측한 자료이다.

이 연구에서는 결측치 제거를 통해 데이터 전처리 작업을 수행하였다. 또한, 월 단위 데이터 자료의 총 30개의 변수 관측치를 가지는 3,291개의 최종 데이터 셋을 추출하였다.

총 30개의 조사된 변수는 위치, 월, 기온, 해면기압, 증기압의 평균, 최대, 최소값과 평균 현지 기압, 이슬점온도, 상대습도의 평균, 최댓값, 풍속의 평균, 최댓값, 강수량, 월합 강수량, 일 최대 강수량, 일사량, 평균 지면온도가 포함된다.

2.1.2 식중독 발생 자료

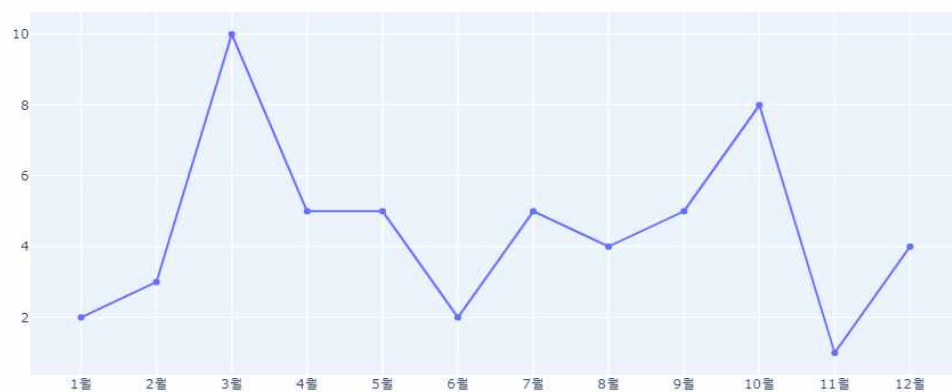
본 연구에 사용된 식중독 발생 자료는 식품의약품안전처가 수집한 전국 식중독 발생 현황 보건의로 빅 데이터 자료 중 2018년 1월부터 12월까지 발생 월, 발생 지역, 원인균 등의 자료가 포함된 식중독 발생 건수 자료를 이용하였다.

본 연구에서는 식중독 발생 현황 빅 데이터 자료에 대해 월별 식중독 사고 기초 통계 분석한 결과 3월~5월, 7월~10월, 12월에 높은 식중독 발병률을 보였다.

필자는 본 전공인 식품영양학과와 위생학 수업을 통해 식중독을 일으키는 미생물도 여러 가지 종류가 있다는 것을 학습하였다. 식중독을 일으키는 미생물은 크게 세균성, 바이러스성으로 분류된다. 세균성 식중독에서도 감염형, 독소형, 독소 감염형 3가지로 분류할 수 있다. 세균성 식중독을 일으키는 원인균 중 감염형 세균은 Salmonella, E.Coli(EPEC, EIEC, ETEC, EHEC, EAEC), Listeria monocytogenes, Campylobacter jejuni, Campylobacter coli, Shigella, Yersinia, Vibrio parahaemolyticus, Vibrio vulnificus으로 총 13종이다. 독소형 세균은 Clostridium botulinum, Staphylococcus aureus으로 2종, 독소감염형 세균은 Clostridium perfringens, Bacillus cereus로 총 3종이다. 바이러스성 식중독을 일으키는 원인균 및 물질은 총 7종으로 Astroviridae, Adenovirus, Norovirus, Hepatitis A형, E형, Rotavirus, Sapovirus가 있다. 바이러스 이 중 Norovirus, Hepatitis A, Rotavirus 가 주 바이러스 식중독 원인 물질이다. 바이러스 물질들은 적은 양으로도 식중독을 일으킬 수 있고 주로 식품과 물을 매개로 하여 감염시킨다.[7]

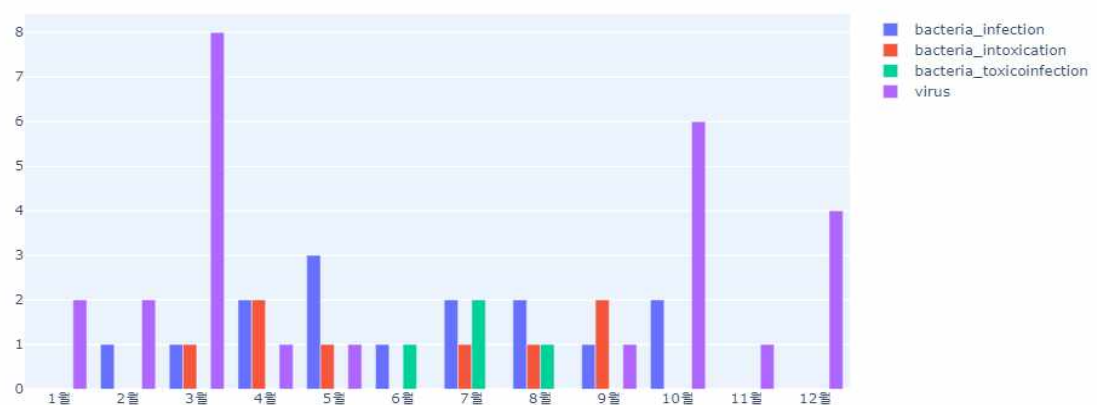
필자는 세균성 식중독 3종, 바이러스 식중독으로 분류하여 기초통계분석을 하였다. 그 결과, 감염형 식중독 세균은 여름, 가을철에 발생률이 높고 독소형 식중독 세균은 봄, 가을에 발생률이 높은 것을 알 수 있다. 독소 감염형 세균은 주로 여름철에 식중독 발병의 원인물질이 되는 것으로 확인되었다. 바이러스성 식중독은 겨울철에 발생률이 높음을 그래프를 통해 알 수 있다. 식중독 발생 원인을 네 가지로 분류하였을 때, 계절적 특징이 두드러지므로 본 연구는 감염형 세균성 식중독, 독소형 세균성 식중독, 독소 감염형 세균성 식중독과 바이러스성 식중독으로 구분하여 기상자료에 따른 식중독 발병률 변이를 확인하였다.

2018 서울시 월별 식중독 발생사고 (단위: 건)



[Fig. 1] 2018 서울시 월별 식중독 발생사고(단위 : 건)

2018년 서울시 월별 식중독 발생사고(단위: 건)



[Fig. 2] 2018년 서울시 월별 식중독 발생사고-원인물질별 (단위: 건)

2.2 분석방법

본 연구의 분석은 각 설명변수들에 대한 랜덤 포레스트(Random Forest)를 사용하여 변수 축약을 하였고, 로지스틱 회귀분석(Logistic Regression)을 사용하여 기상변수에 의한 식중독 발병률을 분석하였다. 분석을 위해 훈련용 데이터(train data)와 평가용 데이터(test data)는 각각 70%와 30%로 설정하여 분석하였다. 또한 빅 데이터 분석에 유용한 R을 사용하여 연구를 진행하였다.

2.2.1 랜덤 포레스트(Random Forest)

랜덤 포레스트는 해당 변수가 상대적으로 얼마만큼 종속변수에 영향을 주는가에 대한 척도의 변수 중요도를 평가할 수 있는 방법 중 하나이다. 변수 중요도는 지니 불순도 감소량이 클수록 데이터 학습 시 중요한 변수로 작용한 것으로 판단할 수 있다.

2.2.2 로지스틱 회귀(Logistic Regression)

로지스틱 회귀는 종속변수가 주로 범주형 데이터인 경우 사용되는 회귀분석법이다. 이는 종속 변수가 0과 1로 구분된다는 것을 의미한다. 본 연구에서 사용된 로지스틱 회귀는 이항분포(Binomial Distribution)를 따르는 이항 로지스틱 회귀(Binomial Logistic Regression)이다.

2.3 변수선택

본 연구에서 식중독 발병률은 식품의약품안전처의 식중독 발생 통계자료의 월별, 지역별, 원인물질별 발생 건수 자료를 바탕으로 한다. 또한 본 연구에서는 식중독 발병률을 범주형 값인 식중독 발생 유무로 정의하였다.

기상 자료에서 구성된 총 30개의 변수들은 종속 변수인 식중독 발병률을 설명하기에 수가 많고 각 변수들의 상관관계가 매우 높기 때문에 변수 축약이 필요하다. 따라서 R의 패키지 중 하나인 'tidymodels'를 활용하여 다중공선성이 나타나는 변수를 제거하고, 척도를 조정해 각 자료를 정규화하고, 가변수를 만드는 작업을 하였다. 그런 다음 랜덤 포레스트를 통해 변수 축약을 진행하였다.

감염형 세균성 식중독 발생 건수 분석에 사용된 총 5개의 설명변수는 최저해면기압, 월합 강수량, 평균풍속, 최대풍속, 평균최저온도이다. 독소형 세균성 식중독 발생 건수 분석에 사용된 7개의 설명 변수는 최저해면기압, 최저상대습도, 일 최다강수량, 일사량, 평균최저온도, 평균풍속, 월 강수량 합이다. 독소 감염형 세균성 식중독 발생 건수에 사용된 7개의 설명변수는 최저해면기압, 최저상대습도, 월합 강수량, 일사량, 최대풍속, 최저기온, 평균풍속이다. 바이러스성 식중독 발생 건수 분석에 사용된 총 8개의 설명변수는 평균 상대습도, 일사량, 일 최다강수량, 최저해면기압, 최저 상대습도, 평균 풍속, 월 강수량 합, 평균 최저기온이다.

3. 연구결과

3.1 로지스틱 회귀 분석 결과

3.1.1 감염형 세균성 원인균에 의한 식중독 발병률 로지스틱 회귀모형

<table 1>과 같이 감염형 세균성 원인균에 의한 식중독 발생률을 종속변수로 로지스틱 회귀분석을 수행한 결과, 최저해면기압, 월합 강수량, 평균풍속, 최대풍속, 평균최저온도가 세균성 원인균에 의한 식중독 발병률에 유의미한 영향을 주는 요인으로 도출되었다. 월합 강수량, 평균풍속, 평균 최저기온이 감염형 세균성 원인균에 의한 식중독 발병률에 정(+)의 영향을 미치고 최저 해면기압, 최대풍속은 부(-)의 영향을 미친다. 다시 말해, 월합 강수량, 평균풍속, 평균최저기온이 높을수록 감염형 세균성 원인균에 의한 식중독 발병률이 높게 나타나며 최저해면기압, 최대풍속이 낮을수록 감염형 세균성 원인균에 의한 식중독 발병률이 높게 나타났다.

<Table 1> result of logistic regression model for infection Bacterial Food Poisoning

	Estimate	Std. Error	z value	Pr(> z)
최저해면기압	-1.19939	1.12669	1.705	0.0881
월합강수량	0.71953	1.00854	-1.065	0.2871
평균 풍속	0.10234	0.87585	0.117	0.9070
최대 풍속	-0.06321	0.97126	0.117	0.9070
평균최저온도	-0.06321	0.97126	-0.065	0.9481

3.1.2. 독소형 세균성 원인균에 의한 식중독 발병률 로지스틱 회귀모형

<table 2>와 같이 독소형 세균성 원인균에 의한 식중독 발병률을 종속변수로 로지스틱 회귀분석을 수행한 결과, 최저해면기압, 최저상대습도, 일 최대강수량, 일사량, 평균최저기온, 평균풍속, 월합 강수량이 독소형 세균성 원인균에 의한 식중독 발병률에 유의미한 영향을 주는 요인으로 도출되었다. 최저 상대습도, 일 최대 강수량, 일사량, 평균최저온도는 식중독 발병률에 정(+)의 영향을 미치고 최저해면기압, 평균풍속, 월 강수량 합은 부(-)의 영향을 미친다. 다시 말해, 최저상대습도, 일 최대 강수량, 일사량, 평균최저기온이 높을수록 독소형 세균성 원인균에 의한 식중독 발병률이 높게 나타나며, 최저해면기압, 평균풍속, 월 강수량 합이 낮을수록 독소형 세균성 원인균에 의한 식중독 발병률이 낮게 나타났다.

<Table 2> Result of Logistic Regression Model for intoxication Bacterial Food
Posoning

	Estimate	Std. Error	z value	Pr(z)
최저 해면기압	-0.3900	1.0463	-0.373	0.709
최저 상대습도	1.0990	1.1615	0.946	0.344
일 최다강수량	0.6514	1.1353	0.574	0.566
일사량	0.8816	1.1645	0.757	0.449
평균 최저온도	0.6468	0.5000	1.294	0.196
평균풍속	-0.1950	0.4261	-0.458	0.647
월합 강수량	-0.2475	0.4956	-0.499	0.617

3.1.3. 독소 감염형 세균성 원인균에 의한 식중독 발병률 로지스틱 회귀모형

<table 3>과 같이 독소 감염형 세균성 원인균에 의한 식중독 발병률을 종속변수로 로지스틱 회귀분석을 수행한 결과, 최저해면기압, 최저상대습도, 월합 강수량, 일사량, 최대풍속, 최저기온, 평균풍속이 독소 감염형 세균성 원인균에 의한 식중독 발병률에 유의미한 영향을 주는 요인으로 도출되었다. 월합 강수량, 일사량, 최저기온은 독소 감염형 세균성 원인균에 의한 식중독 발병률에 정(+)의 영향을 미치고 최저해면기압, 최저상대습도, 최대풍속은 독소 감염형 세균성 원인균에 의한 식중독 발병률에 부(-)의 영향을 미친다. 다시 말해, 월합 강수량, 일사량, 최저기온이 높을수록 독소 감염형 세균성 원인균에 의한 식중독 발병률이 높게 나타나며, 최저해면기압, 최저상대습도, 최대풍속이 낮을수록 독소 감염형 세균성 원인균에 의한 식중독 발병률이 낮게 나타난다.

<Table 3> Result of Logistic Regression Model for Toxicoinfection Bacterial Food
Posoning

	Estimate	Std.Error	z value	Pr(z)
최저 해면기압	-0.57987	0.80293	-0.722	0.470
최저 상대습도	-0.51146	0.85484	-0.598	0.550
월합 강수량	0.34735	0.82716	0.420	0.675
일사량	1.05338	0.96492	1.092	0.275
최대풍속	-0.10808	0.71963	-0.150	0.881
최저기온	2.14662	1.28818	1.666	0.095633
평균풍속	-0.05149	0.53743	-0.096	0.923675

3.1.4 바이러스성 원인균에 의한 식중독 발병률 로지스틱 회귀 모형

<table 2>와 같이 바이러스성 원인균에 의한 식중독 발병률을 종속변수로 로지스틱 회귀분석을 수행한 결과, 평균 상대습도, 일사량, 일 최대 강수량, 최저 해면기압, 최저 상대습도, 평균 풍속, 월 강수량 합, 최저기온이 바이러스성 원인균에 의한 식중독 발병률에 유의미한 영향을 주는 요인으로 도출되었다. 최저 해면기압, 평균 풍속은 바이러스성 원인균에 의한 식중독 발병률에 정(+)의 영향을 미치고 평균 상대습도, 최저 상대습도, 일 최대 강수량, 일사량, 최저기온 부(-)의 영향을 미친다. 다시 말해, 최저 해면기압, 평균풍속이 높을수록 바이러스성 원인균에 의한 식중독 발병률이 높게 나타나며, 평균, 최저 상대습도, 일 최대 강수량, 일사량, 최저기온이 낮을수록 바이러스성 원인균에 의한 식중독 발병률이 높게 나타났다.

<Table 4> Result of Logistic Regression Model for Viral Food Poisoning

	Estimate	Std. Error	z value	Pr(> z)
평균 상대습도	-0.5163	1.1899	-0.434	0.664
일사량	-0.5502	1.1686	-0.471	0.638
일 최대강수량	-1.3452	1.1836	-1.137	0.256
최저 해면기압	0.3580	1.0540	0.340	0.734
최저 상대습도	-0.3866	0.9748	-0.397	0.692
평균풍속	0.2525	1.0375	0.243	0.808
월 강수량 합	-0.1717	0.4484	-0.383	0.702
최저기온	-0.5214	0.4018	-1.298	0.194

4. 고찰 및 결론

본 연구는 기후변화에 따른 식중독 발병률의 상관관계를 확인하기 위하여 식품의약품안전처의 식중독 발생 빅 데이터 자료와 기상청의 기상관측 빅 데이터 자료를 활용하여 로지스틱 회귀 모형을 이용하여 분석하였다. 이 연구는 기존의 “기상요인과 식중독 발병의 연관성에 대한 연구”에 전공 심화적인 내용을 추가하여 진행되었다. 식중독을 발생시키는 원인 중에서 감염형 세균성 원인균, 독소형 세균성 원인균, 독소 감염형 세균성 원인균에 의한 식중독과 바이러스성 원인균에 의한 식중독을 상호 비교하면서 식중독 발병률 변이를 확인하였다. 또한 기상 자료의 변수들의 종류가 많아 변수들의 특성을 고려하여 비교하면서 랜덤포레스트 변수 중요도의 변수 축소과정을 통해서 예측 정확도를 높였다.

본 연구의 결과로 첫 번째, 로지스틱 회귀 분석 결과 세균성과 바이러스 원인균에 의한 식중독 발병률은 기초 통계 분석 시에 확인했던 것과 같이 서로 다른 특성을 가지며, 같은 요인에 의해 영향을 받더라도 세균성 원인균에 의한 식중독 발병률은 정(+)의 영향을 바이러스성 원인균에 의한 식중독 발병률은 부(-)의 영향을 받는 것으로 나타났다. 이는 식중독을 일으키는 세균은 대부분 35°C ~36°C 내외에서 번식 속도가 가장 빠르고 Norovirus, Rotavirus와 같은 바이러스는 낮은 온도에서 살기 좋은 환경이 만들어지기 때문에 식중독을 구분하는 온도

에 따른 원인균의 증식 환경 때문이다.[7] 또한 세균성 원인균 중 감염형 세균성 원인균은 일사량에 의해 영향을 받지 않는 결과가 도출되었지만 독소형과 독소 감염형 세균은 일사량에 의해 영향을 받는 결과를 확인할 수 있었다. 더위는 기온뿐 아니라 습도, 일사량에도 영향을 받는데 같은 기온이라도 습도가 높고 일사량이 강하면 더욱더 내열성이 강한 세균이 증식하기에 좋아하는 환경이 만들어지기 때문이다.[7]

본 연구는 1년 동안 서울에서 일어난 식중독 사고에 대한 신고 또는 보고에 의한 역학 자료이다. 필자는 데이터 마이닝 기술을 이용하여 식품위생 지식과 빅 데이터 기술을 접목하여 어떻게 식중독 사고를 예방할 수 있을지에 대해 고민해 보면서 이 연구를 진행해 보았다. 이 연구를 통해 우리가 식중독에 영향을 주는 요소 중 하나인 기온뿐만 아니라 다양한 기상 요인이 식중독 사고에 영향과 관련이 있다는 것을 직접 확인할 수 있었다. 이 통계 자료를 다룸으로써 기상 요인의 다양한 변수가 식중독 사고의 원인이 될 수 있다는 새로운 지식도 배울 수 있었다. 또한 식품영양 전공 교과목에서 배운 내용과 빅 데이터 분야를 융합해 봄으로써 우리가 살아가고 있는 사회는 모든 데이터 정보가 얹혀있다는 것을 느낄 수 있었다. 이 연구는 전국 단위의 자료가 아닌 특정 지역과 연도의 자료를 사용했기 때문에 이 분석 결과를 일반화하기에는 한계점이 있다. 하지만 필자는 이번 연구를 기반으로 서로 다른 분야인 두 데이터의 연관성과 다루는 법을 익혔다. 후에 심화된 빅 데이터 프로그래밍 기술을 배움으로써 더 큰 단위의 자료를 다루어 볼 것이다. 또한 개인적 위생 관념, 조리 방법도 식중독 사고에 크게 영향을 미치기 때문에 기상요인만으로는 식중독 발생 사고의 영향을 추정하기는 쉽지 않은 면도 있다. 필자는 향후에 이러한 요소들까지 고려하여 식중독 발생 사고를 예측할 수 있는 시스템을 만들어 볼 것이다.

참고 문헌

- [1] 박지애, 김장목, 이호성, 이해진, “기상요인과 식중독 발병의 연관성에 대한 빅 데이터 분석”, *Journal of Digital Convergence*, 2016 Mar, Vol 14, p.319-p.327.
- [2] Xiaoxu Wu, Yongmei Lu, Sen Zhou, Lifan Chen, Bing Xu, “Impact of climate change on human infectious diseases: Emprical evidence and human adaptation”, *Environment International*, Vol.86, p.14-p.23. 2015.
- [3] 신호성, 윤시문, 정진욱, 김정선, “기후변화와 연령특성이 장염 발생에 미치는 영향”, *보건 사회연구*, 2015, p. 186-p. 210.
- [4] R. S. KOVATS, S. J. EDWARDS, S. HAJAT, B.G. ARMSTRONG, K.L. EBI, B.MENNE and The Collaborating Group, “The effect of temperature on food poisoning: a time-series analysis of salmonellosis in ten Europea countries”, *Epidemiol*, Vol.132, p.443-p.453, 2004.
- [5] J.D.Greig, A.Ravel, “Analysis of foodborne outbreak data reported internationally for source attribution”, *International Journal of Food Microbiology*, Vol.130, p.77-p.87, 2009.
- [6] Atul A. Khasnis and Mary D. Nettleman, “Grobal Warming and Infectious Disease”, *Archives of Medical Research*, Vol.36, p689-696, 2005.
- [7] 윤기선, 엄애선, 정윤경, 김경임, 이종경 외 1명, “식품 위생학”, 파워북, p.92-p.95, 2018.