

베이지안 추론을 이용한 전쟁 시뮬레이션과 예측 연구

A Study on the War Simulation and Prediction Using Bayesian Inference

이승용, 유병주, 윤상윤, 방상호, 정재웅
(주)심네트 경영전략본부 R&D 및 V&V연구소

Seung-Lyong Lee(winstar99@simnet.kr), Byung Joo Yoo(bjyoo@simnet.kr),
Sangyoun Youn(syyoun@simnet.kr), Sang-Ho Bang(bsh@simnet.kr),
Jae-Woong Jung(jung7851@simnet.kr)

요약

시간적인 차이를 두고 획득한 이질적인 과거 전쟁 결과 데이터를 하나의 모형으로 구축하는 방법으로 베이지안 추론에 의한 전쟁시뮬레이션 모형을 구축하는 방법을 제안하였다. 과거의 전쟁 결과를 분석하여 미래에 있을 수 있는 전쟁을 예측하는 방법으로 선형회귀모형을 적용하는 방법을 고려할 수 있다. 그러나 역사적으로 시대가 서로 달라 전장 환경의 변화가 반영된 이질적인 두 유형의 자료들이라면 모형의 가정사항 위반으로 하나의 선형회귀모형으로 적합하는 것은 적절하지 않다. 이러한 문제를 해결하기 위해 앞선 시대에 있는 자료를 비정보적 사전분포로 가정하여 사후분포를 구하고 이를 다음 시대에 얻은 자료를 분석하기 위한 사전분포로 활용하여 최종 사후분포를 추론하는 베이지안 추론 방법을 제안하였다. 베이지안 추론 방법의 또 다른 장점은 마코프 체인 몬테 카를로 방법으로 샘플링한 결과를 이용하여 불확실성이 반영된 사후분포나 사후예측분포를 추론할 수 있다는 점이다. 이렇게 했을 때 고전적인 선형회귀모형으로 분석하는 것보다 다양한 정보를 활용할 수 있을 뿐만 아니라 향후 추가적으로 획득되는 자료도 모형에 반영하여 모형을 계속 업데이트시킬 수 있다는 장점이 있다.

■ 중심어 : | 베이지안 추론 | 마코프 체인 몬테 카를로 | 전쟁모형 | 사후분포 | 모델링과 시뮬레이션 |

Abstract

A method of constructing a war simulation based on Bayesian Inference was proposed as a method of constructing heterogeneous historical war data obtained with a time difference into a single model. A method of applying a linear regression model can be considered as a method of predicting future battles by analyzing historical war results. However it is not appropriate for two heterogeneous types of historical data that reflect changes in the battlefield environment due to different times to be suitable as a single linear regression model and violation of the model's assumptions. To resolve these problems a Bayesian inference method was proposed to obtain a post-distribution by assuming the data from the previous era as a non-informative prior distribution and to infer the final posterior distribution by using it as a prior distribution to analyze the data obtained from the next era. Another advantage of the Bayesian inference method is that the results sampled by the Markov Chain Monte Carlo method can be used to infer posterior distribution or posterior predictive distribution reflecting uncertainty. In this way, it has the advantage of not only being able to utilize a variety of information rather than analyzing it with a classical linear regression model, but also continuing to update the model by reflecting additional data obtained in the future.

■ keyword : | Bayesian Inference | Markov Chain Monte Carlo | Combat Model | Posterior Distribution | Modeling & Simulation |

* 본 연구는 행정안전부 재난안전취약핵심역량 도약기술개발사업(시뮬레이션 기반 조직단위 비상대비 훈련기술 개발)의 지원을 받아 수행된 연구임(2018-MOIS33-001)

접수일자 : 2021년 09월 16일
수정일자 : 2021년 10월 20일

심사완료일 : 2021년 10월 20일
교신저자 : 유병주, e-mail : bjyoo@simnet.kr

I. 개요

최근 국방 분야에서는 실제 전쟁에 참여하지 않고 모델링 기술로 과거 전쟁을 재현하거나 시뮬레이션 기술로 간접적으로 전쟁을 체험할 수 있는 체계들이 개발되고 있다. 여기서 모델링은 복잡한 체계나 과정에 관한 관심 분야의 특징적인 부분을 단순화하여 표현하는 과정이며, 시뮬레이션은 모델링한 과정을 시간의 흐름 속에서 반복적으로 재현하는 것이라고 할 수 있다. 이러한 분야에서 무엇보다 중요한 것은 모델링과 시뮬레이션 체계에서 재현되는 콘텐츠의 표현과 내용의 정확성이다. 콘텐츠의 정확성은 정확한 의사소통과 불필요한 자원 낭비를 방지할 수 있는 중요한 요소이기 때문에 특히 고위험을 감수해야 하는 국방 분야에서는 필수적으로 고려해야 할 사항이다.

그래서 전쟁에 대한 모델링이나 시뮬레이션 콘텐츠의 정확도 향상을 위해서 다양한 과거 전쟁 결과를 분석하고 미래에 있을 수 있는 전쟁이나 전투를 보다 정확하게 추론하고자 노력해야 할 것이다. 하지만 이러한 문제를 해결하는 접근방법은 상황별로 너무 다양하여 표준적인 방법을 제시하기는 어려운 것이다. 예를 들어 현재의 청군과 홍군의 전투력을 가지고 전쟁을 했을 때 서로 발생 가능한 피해 규모나 승리 가능성 등에 대하여 예측하는 방법은 통상의 경우 미리 복잡한 전장 환경에 대하여 수학적이거나 공학적으로 모델링한 도구를 이용하여 변화된 환경을 반영한 후 반복적으로 시뮬레이션한 결과를 분석하여 결론을 도출하고자 할 것이다.

컴퓨터 시뮬레이션에 의한 전쟁 결과 예측을 위해서는 그 시뮬레이션을 위한 기초 데이터가 검증된 값으로 입력되어야 하는데 그중에 대표적인 입력 콘텐츠로써 시간에 따른 전투력손실률, 인원 손실률, 장비 손실률, 손실교환비 등이 필요하다. 이러한 자료를 검증하는 방법은 전쟁사 자료 분석, 전투실형, 무기효과 자료 등 다양한 소스로부터 획득한 정확한 데이터를 근거로 반영해야 할 것이다.

이처럼 전쟁 결과를 분석하거나 예측하는 방법은 공학적, 수학적, 통계적 모형 등 다양한 접근방법이 가능할 것이다. 그중 전쟁모형으로서 간단하면서도 오랫동안 활용되는 모델 중의 하나가 Lanchester 모형일 것

이다. 란체스터 모형은 두 부대 간의 상대적 전투력의 관계를 미분방정식으로 표현하고 시간 경과에 따라 발생할 수 있는 피해를 예측할 수 있게 해준다[1].

Morse와 Kimball은 수학적 접근방법으로 조준사격이나 조우전하에서의 란체스터 선형모형을 발전시켰으며, Weiss는 미국의 남북전쟁에 관한 연구를, Peterson은 기갑 전투에서의 란체스터 지수모형 또는 지승모형을 적용하고자 노력하였다[2-4]. 또한, Dietchman과 Schaffer는 비정규전을 위한 란체스터 모형을 제안하였고, Hartley와 Helmbold는 한국전쟁 당시 인천상륙작전 중 미 해병 1사단과 북한군 자료를 활용하여 란체스터 동질 지수모형을 제시하였다[5-7].

한편 이들이 대부분 주어진 자료에 대한 최적화 측면에서 분석을 시도하였다면, 통계적 방법으로 접근한 사례도 있는데 Bracken과 Fricker는 2차 세계대전 중 Ardennes 전역에 대한 자료를 이용하여 란체스터 모형을 발전시켰으며, Wiper는 이 자료에 대한 베이지안 추론을 시도하였다[8-10].

베이지안 추론은 사전분포(Prior)와 우도함수(Likelihood)의 가중 합에 의한 사후분포를 구할 수 있어서 과거의 정보인 사전분포와 최신 정보인 우도함수를 결합하여 사후분포를 구하고, 이를 통해서 사후예측분포를 몬테카를로 적분으로 구해서 미래를 예측할 수 있다. 그래서 과거의 충분한 정보가 없다면 비정보적 사전분포를 부여하고, 과거의 정보가 있다면 그것을 사전분포에 반영하여 모형을 구축함으로써 미래 예측이 가능하다. 특히 베이지안 추론의 경우는 모수들을 하나의 확률변수로 취급함으로써 각각의 모수에 대한 점 추정이 아닌 사후분포 전체를 추정하여 고전적 선형모형보다 미래를 예측하기 위한 많은 정보를 활용할 수 있다.

이러한 장점 때문에 베이지안 추론은 시간의 흐름 속에 역사적인 데이터를 이용해 모형을 업데이트해 나가는데 많이 활용하고 있어서 군사적 관점에서 과거의 전쟁사 자료를 분석하고 예측하는데 적용하기에 적절하다고 판단하였다. 그래서 과거의 전쟁사 자료를 수집하여 미래에 있을 수 있는 전쟁을 예측하고 선형모형으로부터 얻을 수 없는 보다 풍부한 미래 예측 데이터를 활용하는 방법과 사례를 제시하도록 한다. 이를 위해 2장에서는 연구 방법과 모형을 설정하고, 3장에서는 분

석 대상 자료에 대한 베이지안 추론을 다루며, 4장에서는 미래 예측이라는 관점에서 모형이 주는 함의를 해석하고 예측하는 방법을 제시하고자 한다.

II. 연구방법

1. 전쟁모형 관련 기존 연구

군사적 측면에서 전쟁모형은 국가별로 국익과 직결되는 부분이 많아서 대외적으로 발표된 논문은 많지 않다. 기존 연구는 대부분 란체스터 모형에 관한 추가적인 연구, 인공지능경망을 적용한 전쟁모형 연구, 그리고 전쟁사 자료를 이용한 통계적 모형구축 연구들이 있다.

란체스터 모형에 관한 추가적인 연구는 Kress 등에 의해서 전투규칙에 있어서 공자와 방자의 전투력비가 3:1이 되어야 한다는 주제를 확률적 란체스터 모형을 활용하여 논의하였으며[11], 박동현 등은 혼합군에 대한 확률적 란체스터 모형의 정규군사를 통하여 생존병력의 최대화와 전투종료 시간 차이 최소화를 위한 최적화 측면의 접근을 시도하였고[12], 유병주는 란체스터 모형에 대한 통계적 고찰과 해석을 통하여 란체스터모형을 통계적 모형으로 어떻게 변환 적용 가능한지 그 접근방법과 사례를 제시하였다[13].

최근에 인공지능경망을 적용한 전쟁모형 연구는 심홍기 등에 의해서 인공지능경망을 이용한 대대 전투 간 작전 지속능력 예측을 위해 일반 회귀신경망 모형을 적용하여 임무달성 여부를 예측하고자 하였으며[14], 윤성웅 등은 인공지능경망 기반 전술 상황 평가모형을 제안하여 모의 전투훈련 예측에 활용하는 방안을 제시하였다[15].

과거 전쟁사를 통하여 전쟁모형을 연구한 사례는 유병주가 연구한 전쟁사 자료를 이용한 인원손실을 분석 방법론이 있으며, 이는 2차 세계대전의 사단급 전투 자료를 활용하여 포아송 회귀모형, 준포아송 모형, 음이항 모형 등을 적합시켜 분석하는 방법을 비교 분석하였다[16].

기존의 전쟁모형에 관한 연구는 대부분 란체스터 모형에 관한 연구로서 특정 두 전투력 간 전투를 했을 때 시간의 흐름 속에서 서로의 전투력이 어떻게 손실되는지를 추정할 수 있도록 전투력 교환비나 전투력 승수 등에 대한 최적해를 구하고자 하는 방법에 집중되어 있

었다. 반면 최근에는 이러한 경향에서 벗어나 인공지능이나 통계적 접근을 통하여 모형을 새롭게 구축하고 그 모형을 이용하여 미래에 있을 수 있는 전쟁을 예측하기 위한 연구가 활발히 진행되고 있다.

전쟁모형을 개발하기 위해서는 전쟁의 본질을 이해하고 전쟁 양상의 변화를 고려하여 수집된 데이터가 동질의 데이터인지 이질적인지 고려해야 할 필요가 있다. 통계적 모형은 기본적으로 데이터들이 서로 독립적이면서 동일한 분포를 따른다고 가정하고 있어서 전쟁사나 군사적 전투양상의 변화는 그 데이터가 동질의 것이 아니라는 것을 고려해야 하고 더구나 역사적으로 시간의 흐름 속에 얻어진 데이터는 서로 독립적인 동질의 데이터가 아닐 수 있다고 생각해볼 필요가 있다.

선형회귀 분석의 목적은 연속형 자료인 독립변수와 종속변수 간의 관계를 잘 설명할 수 있는 모형을 구축하고, 미래에 있을 수 있는 특정 독립변수에 대한 종속변수의 변화를 예측하고자 함일 것이다. 그러나 고전적인 선형회귀모형의 경우는 미래를 예측할 때 추정된 모수의 고정값을 기준으로 예측하기 때문에 모수 변화에 대한 종속변수의 변화 가능성을 예측하기 어렵다. 반면 베이지안 추론의 경우는 모수도 확률변수로 고려하여 모수의 불확실성을 포함한 사후예측분포 전체를 추론할 수 있다는 장점이 있다[17][18].

2. 고전적 선형회귀모형 적용 방법

분석하고자 하는 데이터가 과거 전쟁이나 전투로부터 획득된 역사적 자료라고 생각해보자. 그러면 역사적 시간의 흐름에 따른 설명하기 어려운 환경변화가 많이 작동되었을 것이라고 볼 수 있다. 그래서 1900년대 초반의 2차 세계대전 자료와 1900년대 후반의 중동전 자료를 서로 독립이면서 동일한 분포를 따르는 자료로 고려한다는 것은 전장 환경의 변화를 생각하면 다소 무리가 있어 보인다. 그래서 역사적 자료 분석을 위해서는 과거의 정보와 현재의 정보를 단계적으로 업데이트할 수 있는 모형이 필요하다. 그리고 고전적 통계분석의 제약 조건으로부터 자유로운 베이지안 추론 방법을 적용하여 적절한 전쟁모형을 구축하고 미래를 예측하기 위한 분석을 수행하는 것은 매우 의미 있는 도전일 것이다.

그래서 연속형 변수인 자료에 대하여 변수 간의 관계가 원인과 결과의 관계를 맺고 있다면 일반적으로 선형회귀 모형을 구축하여 분석할 것이 권장되어 진다. 그래서 우리가 분석하고자 하는 자료는 개전시 전투력비가 종전시 전투력비에 어떠한 결과로 귀결되는지 선형 모형으로 적합시킬 수 있는데 전투력비의 자료 특성상 어떤 비율을 나타내고 있으므로 로그변환을 통하여 모형을 구축할 수 있다. 이러한 점을 고려하여 원자료에 대한 모형 $E = \exp(\beta_0)S^{\beta_1}$ 에 대하여 Faraway가 제시한 방법으로 $y = \log(E)$, $x = \log(S)$ 로 변환하면 고전적인 선형회귀모형과 같아지며 그 모형은 아래와 같다[19]. 여기서 과거 자료를 Data1이라고 하고 최근 자료를 Data2라고 했을 때 이것을 하나의 데이터로 모은 것을 Data라고 하면 이에 대하여 각각 선형회귀모형을 적용한다면 아래와 같은 모형이 될 것이다.

$$\text{Data1 모형(C1): } y_{i1} = \beta_{01} + \beta_{11}x_{i1} + \varepsilon_{i1} \quad (1)$$

$$\text{Data2 모형(C2): } y_{i2} = \beta_{02} + \beta_{12}x_{i2} + \varepsilon_{i2} \quad (2)$$

$$\text{Data 통합모형(C) : } y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + \varepsilon_{ij} \quad (3)$$

여기서 $i = 1, \dots, n$ 이고 $j = 1, 2$ 로서 Data1과 Data2에 대한 인덱스라고 생각하자. 이때 β 에 대한 최대우도 추정량을 행렬의 형태로 표현하면 아래와 같다.

$$\hat{\beta} = (X'X)^{-1}X'y \quad (4)$$

여기서 X 는 변수가 하나인 경우 $n \times 2$ 의 디자인 행렬이며 X' 은 독립변수 X 의 전치행렬이다. 그래서 이 모형에서 모수 $\beta = (\beta_0, \beta_1)$ 에 대한 최대우도 추정량은 위 식(4)와 같다.

3. 베이지안 추론 방법

한편 Gilks 등(1996)과 Gelman 등(2004)에 의하면 베이지안 선형회귀 모형의 경우 일반적으로 사전정보가 없는 경우에 우도함수 (5)에 대하여 모수들 즉 β 와 σ^2 에 대하여 비정보적 사전분포 (6)로 가정할 수 있다 [17][18].

$$y | \beta, \sigma^2, X \sim N(X\beta, \sigma^2 I) \quad (5)$$

$$p(\beta, \sigma^2) \propto \frac{1}{\sigma^2} \quad (6)$$

비정보적 사전분포와 정규분포의 우도함수를 가정하는 경우에 β 와 σ^2 에 대한 사후분포는 아래와 같다.

$$\beta | y, \sigma^2 \sim N((X'X)^{-1}X'y, \sigma^2(X'X)^{-1}) \quad (7)$$

$$\sigma^2 | y \sim IG\left(\frac{n-p}{2}, \frac{1}{2}(y - X\beta)'(y - X\beta)\right) \quad (8)$$

여기서 식(7)과 (8)은 완전조건부 분포의 형태이기 때문에 Gibbs 샘플러를 이용하여 Markov Chain Monte Carlo(MCMC) 방법으로 사후분포를 샘플링할 수 있다[17][18]. 그래서 베이지안 선형회귀 모형의 모수 분포를 시뮬레이션하는 방법은 첫 번째로 식(8)에 의해서 새로운 σ^{2*} 을 샘플링하고 그 σ^{2*} 를 식(7)의 σ^2 에 대체하여 넣고 이변량 정규분포로부터 새로운 β^* 들을 샘플링하는 절차로 모수들의 분포를 구할 수 있다.

MCMC 방법에 의한 베이지안 추론은 특정 결합분포를 추정하기 위해 샘플링이 가능한 조건부분포나 주변 분포, 또는 이용 가능한 분포에서 마코프 연쇄를 이용하여 표본을 충분히 추출하여 몬테카를로 방법으로 적분함으로써 결합분포의 기대치나 분산 또는 분포 전체를 추론하는 방법이다. 이러한 추론의 특징은 수치상으로 계산하기 어렵거나 샘플링이 불가능한 분포에 대해서도 추론이 가능할 뿐만 아니라 사후분포를 과거 정보인 사전분포와 현재 정보인 우도함수의 가중 합 형태로 추론하기 때문에 점진적인 정보의 업데이트가 가능하다는 특징이 있다.

베이지안 선형회귀 모형에서 사전분포에 대한 정보가 있는 경우 통상 Normal-Inverse Gamma (NIG) 분포로 고려할 수 있으며 결합분포를 $NIG(m_0, V_0, a_0, b_0)$ 로 정의하면 아래의 식(7)과 식(8)의 곱의 형태로 정의된다.

$$\sigma^2 \sim IG(a_0, b_0) \quad (9)$$

$$\beta | \sigma^2 \sim N(m_0, \sigma^2 V_0) \quad (10)$$

이때 사후분포를 $NIG(m_1, V_1, a_1, b_1)$ 으로 정의하면 세부적인 모수의 정의는 아래와 같다.

$$m_1 = V_1(V_0^{-1}m_0 + X'y), \quad V_1 = (X'X + V_0^{-1})^{-1},$$

$$a_1 = a_0 + \frac{n}{2},$$

$$b_1 = b_0 + \frac{1}{2}(y'y + m_0'V_0^{-1}m_0 - m_1'V_1^{-1}m_1)$$

그래서 β 와 σ^2 는 아래와 같은 사후분포를 갖는다.

$$\beta \mid \sigma^2, y \sim N(m_1, \sigma^2 V_1) \quad (11)$$

$$\sigma^2 \mid y \sim IG(a_1, b_1) \quad (12)$$

종합적으로 분석하고자 하는 전쟁모형은 비정보적 사전분포인 경우는 Gibbs 샘플러를 이용하여 MCMC 방법으로 식(7)과 (8)을 이용해서 사후분포를 샘플링할 수 있으며, NIG분포를 사전분포로 가정한 경우 사후분포는 식(11)과 (12)에서 샘플링이 가능하다.

III. 전쟁모형의 베이지안 추론

1. 분석 대상 자료 소개

Dupuy(1987)의 Understanding War를 참조하면 1939년부터 1945년까지 있었던 2차 세계대전 당시 94개의 전쟁결과(Data1)와 1967년부터 1973년까지의 아랍동맹과 이스라엘의 중동전 전쟁결과 52개의 자료(Data2)를 제공하고 있다[20]. 여기에는 개전시 공자와 방자의 전투력비(S)와 종전시 공자와 방자의 전투력비(E)를 도표 형식으로 제공하고 있는데 정리하여 기초 통계량을 산출하면 아래와 같다[표 1].

표 1. 전쟁사 자료 기술 통계량 요약

구 분	관찰치(N)	개전시 전투력비(S)	종전시 전투력비(E)
Data1	94	1.76	1.28
Data2	52	1.33	1.95

두 가지의 자료를 비교하여 정리해보면 동질적이면 서로 이질적인 자료라고 할 수 있다. 두 가지 모두 동일한 전투력 계산방법으로 계산했다는 측면에서 동질적인 자료라고 볼 수 있으나 2차 세계대전 시 자료와 이후 시기적으로 20년 이상이 경과한 시점의 중동전 자료라는 점을 고려한다면 이질적인 자료라고 할 수 있다. 특히 20년 이상의 시간 경과에는 무기체계의 변화나 전쟁 양상의 변화가 충분히 있을 수 있으므로 과거의 자료와 동일한 성격의 자료라고 보기 어려운 부분이 있다.

2. 선형회귀모형을 이용한 전쟁모형 구축

두 가지 자료의 특성을 고려하여 146개의 모든 자료를 하나의 모집단으로 고려하여 식(3)을 이용하여 분석하는 모형(C), 두 자료가 서로 이질적인 것으로 고려하여 Data1 자료만을 이용하여 식(1)로 적합시킨 모형(C1), Data2 자료를 고려하여 식(2)을 적용한 모형(C2)으로 구분하여 분석하면 [표 2]와 같은 추정 결과를 얻을 수 있다.

표 2. 고전적 선형회귀모형 추정 결과

모 형	$\hat{\beta}_0$			$\hat{\beta}_1$		
	추정치	S.E.	p-value	추정치	S.E.	p-value
C1	0.04	0.03	<0.01	0.39	0.04	<0.01
C2	0.44	0.11	0.24	0.36	0.12	<0.01
C	0.19	0.04	<0.01	0.34	0.05	<0.01

모형들의 P-value를 고려하면 전반적인 수준에서 잘 적합된 것으로 평가할 수 있다.

고전적 선형회귀모형으로 적합하여 모수를 추정할 결과 예상했던 바와 같이 C1과 C2 모형이 다소 이질적이라는 부분을 확인할 수 있다. [그림 1]을 참조하면 실제 자료의 퍼짐이나 군집도 두 자료가 다소 이질적인 것으로 보인다. 그리고 모형 적합시 표준화 잔차가 2보다 큰 Outlier는 C1 모형의 경우는 4개(+표시), C2 모형에서는 3개(x표시)가 식별되어 적절한 조치가 필요한 것으로 판단된다. 이러한 것을 고려하여 모형을 수정 보완해 가는 과정이 필요하나 여기서는 생략하기로 한다.



그림 1. 로그선형모형 적합 결과

사실 우리가 고려하고 있는 두 자료는 이질적인 자료이지만 고전적 선형회귀모형으로 접근하고자 한다면 하나의 모집단으로 고려하여 분석하는 C모형으로 분석할 수밖에 없다. 그리고 세밀한 부분을 수정 보완하고 모형을 검증하여 완성해나가는 절차로 분석해야 할 것이다. 그래서 4개의 비정상 관찰치에 대한 문제 해결 방법은 분석 대상 자료에서 제거하고 모형을 적합하거나 다른 모형을 적용하여 분석할 필요가 있다.

한편 데이터의 특성을 고려하여 두 개의 모형을 분리하여 구축한 모형 C1, C2의 경우 과연 현재 시점에서 앞으로 있을 수 있는 상황을 예측하기 위해 어떠한 모형을 적용할 것인가 고민스러운 선택을 해야 할 것이다. 이 경우 모든 데이터를 활용한 모형이 아닌 부분적 데이터를 사용하게 됨으로써 정보의 손실이 발생할 것이다.

3. 베이지안 추론 방법에 의한 전쟁모형 구축

베이지안 추론 방법을 적용한 전쟁모형은 네 가지 경우를 고려하였으며 그 추론 결과는 [표 3]과 같다. 주어진 Data1과 Data2에 대하여 각각 식(7), (8)을 이용하여 비정보적 사전분포를 반영한 베이지안 모형을 B1, B2 모형으로, 그리고 모든 데이터를 하나의 모집단으로 고려하여 식(7)과 (8)처럼 비정보적 사전분포를 고려한 베이지안 선형회귀 모형을 구축한 것을 BC 모형, Data1에 대하여 우선적으로 베이지안 선형회귀 모형을 식(7), (8)을 적용하여 얻은 사후분포를 Data2에 대한 사전분포로 간주하여 식(11), (12)로 부터 얻은 사후

분포 모형을 B 모형으로 고려하였다. 샘플링은 총 21,000개를 하여 1,000개는 Burn-In으로 버리고 추정 한 결과 각 모형에 대한 사후분포는 아래와 같다[표 3].

표 3. 베이지안 방법에 의한 전쟁모형 추론 결과

모형	$\hat{\beta}_0$		$\hat{\beta}_1$		$\hat{\sigma}^2$	
	mean	SD	mean	SD	mean	SD
B1	0.04	0.03	0.39	0.04	0.11	0.02
B2	0.44	0.11	0.36	0.13	0.60	0.13
BC	0.19	0.05	0.34	0.05	0.31	0.04
B	0.11	0.03	0.38	0.04	0.31	0.04

이렇게 최종적으로 얻어진 BC 베이지안 선형회귀 모형의 사후분포는 Data1과 Data2의 모든 자료를 활용한 결과가 된다. 한편 B 모형의 경우는 Data1로부터 얻은 정보를 사전분포로 하여 Data2의 정보를 업데이트시키는 형태의 모형이며, 이 모형의 각 모수의 사후 분포에서 주요 분위수는 [표 4]와 같다.

베이지안 추론 결과를 고전적 접근방법과 비교하기 위해 [표 2] C 모형의 $\hat{\beta}_0$ 추정값 0.19와 $\hat{\beta}_1$ 추정값 0.34는 [표 4]에 있는 95% 신뢰구간인 $\hat{\beta}_0=(-0.01, 0.09)$, $\hat{\beta}_1=(0.34, 0.45)$ 와 비교하면 구간 밖에 위치하여 동일한 분포라고 판단하기 어렵고 베이지안 모형보다 과대 과소 추정된 것으로 볼 수 있다.

표 4. 베이지안 B모형의 사후분포 추론 결과 요약

모수	2.5%	25%	50%	75%	97.5%
$\hat{\beta}_0$	-0.01	0.02	0.04	0.06	0.09
$\hat{\beta}_1$	0.34	0.37	0.39	0.41	0.45
$\hat{\sigma}^2$	0.08	0.09	0.10	0.11	0.12

C 모형과 베이지안 추론에 의한 B1, B2, B 모형의 추정 결과의 차이를 좀 더 쉽게 설명하기 위해 베이지안 추론에 의한 $\hat{\beta}_0$ 의 사후분포를 도식화하면 [그림 2]와 같다.

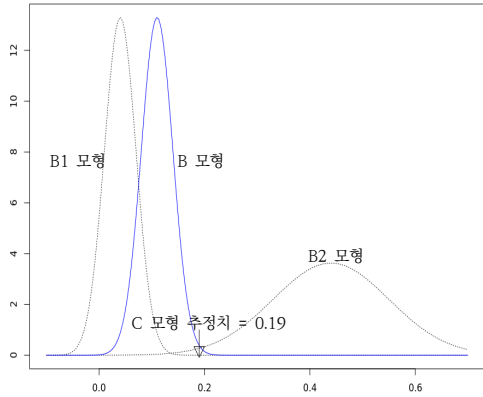


그림 2. 베이지안 모형 $\hat{\beta}_0$ 사후분포와 C모형의 추정치 비교

[그림 2]의 베이지안 사후분포 B1 모형의 경우는 정규분포 $N(0.04, 0.03^2)$ 을 따르고 B2 모형의 경우는 $N(0.44, 0.11^2)$ 의 분포를 따르며, B 모형은 $N(0.11, 0.03^2)$ 의 분포를 따르는 분포로 추정된 것을 확인할 수 있다. 이 그래프가 시사하는 바는 첫 번째로 베이지안 추론 결과 Data1의 $\hat{\beta}_0$ 분포는 Data2의 $\hat{\beta}_0$ 분포와 명확히 구분되며 Data1 추정결과를 사전분포로 하여 Data2에 대한 베이지안 추론결과 B 모형의 $\hat{\beta}_0$ 분포는 B1 모형과 B2 모형의 가중 합인 것을 알 수 있다. 그리고 C 모형의 $\hat{\beta}_0$ 의 추정치 0.19는 B 모형의 사후분포의 95% 신뢰구간에서 벗어나 두 모형의 추정 결과는 유사하다고 말할 수 없다.

IV. 전쟁모형의 해석과 미래 예측

1. 전쟁모형의 해석

여기서 사용된 자료들은 모두 공자의 전투력을 방자의 전투력으로 나눈 비율 값이다. 이런 점을 고려하여 개전시 전투력비와 종전시 전투력비의 관계를 분석하기 위해 모형을 구축했고, 구축된 모형을 기준으로 미래를 예측해보는 것은 매우 흥미로운 일이다. 그래서 전쟁 이전에 어느 정도의 상대적 전투력비를 가지고 전쟁에 대비해야 할 것인가라는 문제도 그 중에 하나일 것이다. 2차 세계대전의 결과만을 반영한 C1 모형의 추정 결과는 $E = \exp(0.04) * S^{0.39}$ 이고 중동전 자료를

적합시킨 C2 모형의 추정 결과는 $E = \exp(0.44) * S^{0.36}$ 으로 개전시와 종전시 전투력비를 산출하면 아래와 같다.

표 5. 개전시 전투력비별 종전시 전투력비와 전투효율성 추정

전투력비/전투효율성	S = 1	S = 2	S = 3
C1 모형	1.04 / 1.04	1.36 / 0.68	1.60 / 0.53
C2 모형	1.55 / 1.55	2.00 / 1.00	2.31 / 0.77

2차 세계대전 당시의 자료인 C1 모형과 중동전 자료를 활용한 C2 모형의 종전시 전투력비는 전반적으로 향상되어 효율성이 향상된 것을 알 수 있으며, 이는 분명히 전쟁양상의 변화가 있었다는 것을 방증할 수 있는 하나의 자료가 될 수 있다. 한편 2차 세계대전 당시의 자료를 기준으로 분석한 C1 모형에 의하면 공자는 동일한 전투력으로 공격하는 것이 2배나 3배의 전투력으로 공격하는 것보다 효율성이 좋다. 3배의 전투력으로 공격을 개시한다면 종전시에는 3배 이상의 전투력으로 종료되어야 하지만 여기서는 방자보다 더 많은 피해를 본 1.6배의 전투력으로 종전이 된다. 반면 중동전 자료인 C2 모형의 경우 2배의 전투력으로 시작시 2배로 종전되어 전투력 운용의 효율성이 다소 개선되었다고 할 수 있다. 이러한 현상들은 공자의 경우 전투력 운용 특성상 방자보다는 많은 위협에 노출되어 더 많은 피해를 보기 때문이라고 판단되며, 2차 세계대전 당시보다는 중동전시 무기체계의 개선으로 공자의 기습 효과 등으로 전투력 운용의 효율성이 개선된 것이 반영되었을 것으로 해석된다.

2. 전쟁모형을 이용한 미래 예측

베이지안 방법에 의한 전쟁모형 추론은 모수 추정의 불확실성까지 반영한 사후예측분포를 시뮬레이션으로 생성하여 분석할 수 있다. B모형에 대한 사후예측분포를 생성하여 초기 방자와 동일한 전투력으로 공격한 결과(유형 #1)와 초기 3배의 전투력으로 공격한 결과(유형 #2)를 도식하면 아래와 같다.

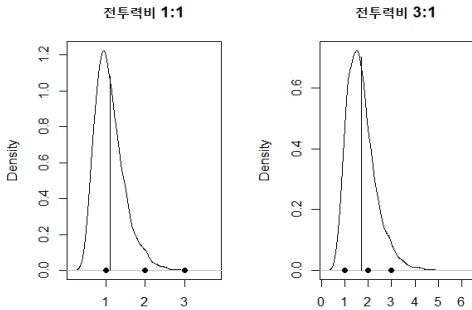


그림 3. 개선시 전투력비별 사후예측분포 추론 결과

[그림 3]에서 수직 실선으로 분포의 평균 위치를 표시하였고, 종전시 전투력비가 1, 2, 3이 되는 위치를 점으로 표시하였다. 그리고 3:1 이상의 전투력으로 종전될 확률은 몬테칼로 시뮬레이션 결과가 3보다 큰 값의 비율(즉, 그래프에서 3보다 큰 지역의 분포를 적분)을 계산하면 된다.

표 6. 종전시 전투력비 수준별 승리 확률

전투력비	평균	1:1 이상	2:1 이상	3:1 이상
유형 #1	1.10	55.1%	2.7%	<0.01%
유형 #2	1.71	90.9%	26.9%	4.0%

위에서 도출된 결과를 보면 방자와 동일한 전투력비로 공격시 2배의 전투력으로 승리할 확률은 2.7%이며 3배의 전투력으로 공격시 2배로 이길 확률은 26.9%가 된다. 이와 같이 베이지안 추론은 몬테칼로 방법으로 사후예측분포 전체를 추론할 수 있는 장점이 있어서 [표 6]과 같은 확률을 구할 수 있다.

검증 차원에서 고전적 접근방법인 [표 5]와 비교하면 유형 #1의 결과 평균이 모형 C1과 C2 값 (1.04, 1.55)의 중간에, 유형 #2의 결과는 모형 C1과 C2 값 (1.60, 2.31)의 중간에 위치하여 합리적인 결과가 도출되었음을 확인할 수 있다.

V. 결론

시간적인 차이를 두고 획득하여 전장 환경의 변화가 반영된 이질적인 전쟁결과 데이터를 다루는 방법으로

베이지안 추론에 의한 전쟁모형을 구축하는 방법을 제안하였다. 시간의 차이는 무기체계나 전쟁 양상의 변화가 있을 수 있으며 이는 데이터의 분포가 동일하지 않고 이질적일 수 있다는 점을 고려해야 하며 그래서 단순히 모든 데이터를 통합하여 하나의 선형회귀모형으로 접근한다면 추정치를 왜곡하여 분석할 수 있다는 점을 설명하고 사례를 보였다.

그리고 전쟁 결과 자료를 수집 분석하면서 고전적인 선형회귀모형을 구축하여 분석하는 방법보다 MCMC 방법을 이용한 베이지안 추론이 모수의 사후분포 전체를 추론할 수 있어서 다양한 정보를 활용할 수 있는 장점에 관하여 사례를 보이고 설명하였다. 그뿐만 아니라 베이지안 추론은 전쟁모형의 해석과 미래 예측에서도 모수의 불확실성이 포함된 사후예측분포를 추론할 수 있어서 다양한 예측이 가능함을 확인하고 설명하였다.

최근 국방 분야에서 실전적 분석과 훈련 방법의 대안으로 대두되고 있는 모델링 및 시뮬레이션 체계에서 재현되는 콘텐츠의 표현과 내용 정확성은 단순히 정확한 결과의 도출을 넘어서 고위험을 수반하는 국방 분야의 효율성을 기하는 필수적인 요소이다. 그래서 어떠한 모형을 왜곡해서 적용하여 발생할 수 있는 오류나 사회적 비용을 줄이고 보다 풍부하고 정확한 콘텐츠를 제공할 수 있는 베이지안 추론 방법의 제안은 국방의 효율화에 기여할 수 있는 방안이라고 볼 수 있다.

그리고 이렇게 개발된 모형은 국가적인 차원에서 과거 전쟁 결과 자료들을 장기간 지속해서 수집하여 양적으로 많은 데이터를 적용한다면 모형의 정밀도는 더욱 향상되며, 콘텐츠의 신뢰성이 더 높아질 것이다. 그래서 이러한 모델링 결과를 활용하여 국가의 총력전 차원의 대비, 비상 기획 업무, 미래 발생 가능한 전쟁 대비 등의 분야에 적용한다면 그 효과를 극대화할 수 있을 것이다.

한편 이번 연구는 Dupuy(1987)의 저서에서 제공한 전투력 데이터를 근거로 연구하다 보니 그 전투력 산출 방법의 합리성과 정확성을 확인할 수 없었고, 더욱이 책자에서 제공하는 전투결과 외의 세부적인 내용은 확인이 제한되어 기초 자료의 정확성을 검증할 수 없었다는 한계가 있다.

참고 문헌

[1] F. W. Lanchester, *Aircraft in Warfare: The Dawn of the Fourth Arm*, Constable and Company, London, 1916.

[2] P. M. Morse and G. E. Kimball, *Methods of Operation Research*, Wiley, New York, 1951.

[3] H. K. Weiss, "Combat Models and Historical Data: The U.S. Civil War," *Operations Research*, Vol.14, No.5, pp.759-790, 1966.

[4] R. H. Peterson, "On the Logarithmic Law of Attrition and Its Application to Tank Combat," *Operations Research*, Vol.15, pp.557-558, 1967.

[5] S. J. Dietchman, "A Lanchester Model of Guerrilla Warfare," *Operations Research*, Vol.10, pp.818-827, 1962.

[6] M. B. Schaffer, "Lanchester Models of Guerrilla Engagements," *Operations Research*, Vol.16, pp.457-488, 1968.

[7] D. S. Hartley and R. L. Helmbold, "Validating Lanchester's Square Law and Other Attrition Models", *Naval Research Logistics*, Vol.42, pp.609-633, 1995.

[8] J. Bracken, "Lanchester Models of Ardennes Campaign," *Naval Research Logistics*, Vol.42, pp.559-577, 1995.

[9] R. D. Fricker, "Attrition Models of the Ardennes Campaign," *Naval Research Logistics*, Vol.45, pp.1-22, 1998.

[10] M. P. Wiper, L. I. Pettit, and K. D. S. Young, "Bayesian Inference for a Lanchester Type Combat Model," *Naval Research Logistics*, Vol.47, pp.541-558, 2000.

[11] M. Kress and I. Talmor, "A New Look at The 3 : 1 Rule of Combat through Markov Stochastic Lanchester Models," *J. of the Operational Research Society*, Vol.50, No.7, pp.733-744, 1999.

[12] D. Park, D. Kim, H. Moon, and H. Shin, "Gaussian Approximation of Stochastic Lanchester Model for Heterogeneous Forces," *J. of Korean Institute of Industrial Engineers*,

Vol.42, No.2, pp.86-95, 2016.

[13] B. J. Yoo, "Statistical review and explanation for Lanchester model," *The Korean Journal of Applied Statistics*, Vol.33, No.3, pp.335-345, 2020.

[14] 심흥기, 김승권, "인공신경망을 이용한 대대전투간 작전지속능력 예측," *지능정보연구*, 제14권, 제3호, pp.25-39, 2008.

[15] S. w. Yoon and S. H. Lee, "ANN-based Evaluation Model of Combat Situation to Predict the Progress of Simulated Combat Training," *한국컴퓨터정보학회논문지*, Vol.22, No.7, pp.31-37, 2017.

[16] 유병주, "전쟁사 자료를 이용한 인원손실률 분석 방법론," *한국국방경영분석학회지*, 제46권, 제1호, pp.31-42, 2020.

[17] Andrew Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman and Hall, 2004.

[18] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman & Hall/CRC, 1996.

[19] J. Julian Faraway, *Linear Models with R*, Chapman and Hall, London, 2015.

[20] T. N. Dupuy, *Understanding War: History and Theory of Combat*, Paragon House, 1987.

저자 소개

이 승 용(Seung-Lyong Lee)

정회원



전략본부 본부장

〈관심분야〉 : 국방 M&S, 국가위기관리, 네트워크, IT 등

- 1982년 2월 : 육군사관학교 공학과 (이학사)
- 1997년 2월 : 동국대학교 외교안보 학과(석사)
- 2017년 2월 : 아주대학교 NCW학과(공학박사)
- 2016년 3월 ~ 현재 : (주)심네트 경영

유 병 주(Byung Joo Yoo)

정회원



- 1988년 2월 : 육군사관학교 건축공학(이학사)
- 2001년 2월 : 고려대학교 통계학과(이학석사)
- 2011년 8월 : 고려대학교 통계학과(통계학박사)
- 2021년 7월 ~ 현재 : (주)심네트 경영

전략본부 책임연구원

〈관심분야〉 : 빅데이터분석, 국방M&S, 테이터사이언스 등

정 재 응(Jae-Woong Jung)

정회원



- 1996년 2월 : 우석대학교 회계학과(학사)
- 2014년 2월 : 조선대학교 정책대학원 국방정책학과(국방학석사)
- 2013년 12월 ~ 2017년 12월 : 제9보병사단/30기계화보병사단 작전/정보과장

▪ 2020년 3월 ~ 현재 : (주)심네트 R&D 및 V&V 연구소 책임연구원

〈관심분야〉 : IT, 국방, 연동

윤 상 윤(Sangyoun Youn)

정회원



- 1975년 2월 : 육군사관학교 병기공학(학사)
- 1983년 2월 : 고려대학교 산업공학과(공학석사)
- 1990년 8월 : 영국 맨체스터대학 산업공학과(산업공학박사)
- 2021년 2월 ~ 현재 : (주)심네트 경영

전략본부 연구소장/책임연구원

〈관심분야〉 : 국방M&S, 분석평가, 재난안전 등

방 상 호(Sang-Ho Bang)

정회원



- 1997년 2월 : 부산외국어대학교 인도학과(학사)
- 2013년 2월 : 한성대학교 국방과학대학원 국방M&S학과(공학석사)
- 2009년 11월 ~ 2014년 11월 : 한미연합사령부/합동참모본부 지상전 분석장교

▪ 2019년 9월 ~ 현재 : (주)심네트 R&D 및 V&V 연구소 책임연구원

〈관심분야〉 : IT, 국방, 연동