

Bayesian Analysis for Uncertainty of Radiocarbon Dating

Youngseon Lee^{a,1} · Jaeyong Lee^a · Jangsuk Kim^b

^aDepartment of Statistics, Seoul National University

^bDepartment of Archaeology and Art History, Seoul National University

(Received December 15, 2014; Revised February 16, 2015; Accepted February 16, 2015)

Abstract

Use of radiocarbon dating is increasing for chronology; however, its variability and discrepancy with existing chronologies can cause doubts in regards to credibility. In this paper, we explore factors that influence radiocarbon dating variabilities. We obtained estimated radiocarbon ages by sending identical samples to several labs multiple times. A Bayesian method was used to analyze the obtained data. From the analysis, we conclude that some factors (such as type of labs and megasamples) can induce variability when estimating radiocarbon age. We identify the size of variability caused by each factor and analyze the estimated variability in each lab corresponds with the reported variability.

Keywords: radiocarbon dating, archaeological chronology, Bayesian statistics, consensus value, variability

1. 서론

미국의 리비(Willard F. Libby)에 의해 1940년대에 개발된 방사성탄소연대측정법은 고고학 연대측정의 한 방법으로 널리 쓰이고 있다. 한국고고학에서는 1990년대 후반부터 이용이 급증하여 현재는 9000건이 넘는 연대측정결과가 보고되어 있다. 하지만, 한국고고학에서는 방사성탄소연대측정치가 연대판정에 적극적으로 이용되지 못할 뿐 아니라 방사성탄소연대 회의론 또는 무용론까지도 제기되고 있다. 그 이유는, 첫째, 동일한 시점을 나타낼 것으로 기대되는 시료의 측정결과들이 고고학적 동시기로 보기 어려울 만큼 서로 간에 큰 차이를 보이기도 하고, 둘째, 탄소연대측정결과가 고고학적으로 이미 수립된 기존의 편년과 현저한 차이를 보이는 경우가 흔히 발생하기 때문이다. 탄소연대측정결과에 대한 부정적 경향은 편년에 문헌사료를 적극적으로 이용하는 삼국시대의 고고학에서 가장 두드러지는데, 이는, 위의 두 문제가 설사 해결된다고 하더라도 탄소연대측정은 그 결과가 원천적으로 가질 수밖에 없는 수십 년 또는 1세기의 오차로 인해 역사적 사건과 고고학 자료의 세밀한 시간적 배열에 적절한 답을 제공하지 못할 것이라는 인식에 상당부분 기인한다. 이로 인해, 방사성탄소연대측정결과를 어떻게 수용하여야 할 것인지는 한국고고학계에서 논란에 휩싸여 있다 (Kim, 2014).

This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2013S1A5B6043901).

¹Corresponding author: Department of Statistics, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 151-747, Korea. E-mail: lysstat@gmail.com

이러한 논란은 비단 한국고고학계만의 문제는 아닌데, 토기의 양식변천과 문헌사료를 고고학 자료의 연대판정에 적극적으로 이용하는 방법론적 전통을 가진 일본, 중국, 그리스, 이탈리아 등 구대륙의 고고학에서도 유사하게 나타난다. 각지에서 공통적으로 발견되는 논란의 초점에 놓여 있는 것은 방사성탄소연대측정결과의 신뢰성이 어느 정도인지, 변동성이 어디에서 기인하는 것인지, 그리고 이를 어느 정도까지 고고학적으로 용인할 수 있는지의 문제이다. 따라서, 방사성탄소연대측정법은 현재까지 개발된 연대측정법 중 가장 많이 이용되고 있고 고고학연구에 지극히 중요한 파급효과를 가져왔음에도 불구하고, 실제 적용에 있어서는 불확실성 제거와 변동성에 대한 이해라는 과제 역시 동시에 안고 있다.

이에, 1988년 13회 국제방사성탄소 학회에서는 방사성탄소연대측정법의 신뢰성을 점검해야 할 필요성을 인식하고, 국제 방사성탄소 비교(International Radiocarbon Intercomparison; IRI)라는 이름 하에 주기적으로 세계각지의 탄소연대측정기관들이 동일한 샘플을 대상으로 탄소연대를 측정하고 이를 보고하는 실험이 이루어져 왔다. 이 비교실험의 가장 큰 목적은 측정기관 간에 측정결과의 차이가 있는지를 점검하는 것으로서, 지금까지 5차에 걸쳐 이루어졌으며, 6차 실험이 예정되어 있다.

Rozanski는 IRI 실험의 결과로 얻어진 자료를 바탕으로 하여, 하나의 샘플에서 얻은 여러 탄소연대들로부터 하나의 합의값(Consensus value)을 계산하는 방법을 제안했다 (Rozanski 등, 1992). 또한 기관마다 측정한 탄소연대에 차이가 있는 것을 바탕으로, 어떤 요인이 이러한 차이를 유발하는지에 관한 검증도 시행했다. 그러나 이러한 연구는, 기관 간에 존재하는 변동성을 이해하기 위한 연구는 아니었으며 단지 샘플의 탄소연대합의값을 효과적으로 계산하기 위한 방법을 찾으려는 연구의 일부로서 시행되었다. Scott는 3차 국제 방사성탄소 비교(the Third International Radiocarbon Intercomparison; TIRI)와 4차 국제 방사성탄소 비교(the Fourth International Radiocarbon Intercomparison; FIRI) 실험결과로부터 각 측정기관이 보고하는 변동성이 실제 측정치들로부터 계산한 변동성보다 과소 추정되어 있다는 것을 발견하고, 이를 보정할 수 있는 오차보정치(Error multiplier)를 제안하게 된다 (Scott, 2003). 이 연구에서도 측정치에 차이를 주는 변동성에 관심을 가지기는 했으나 이러한 변동성이 야기되는 요인과 변동성의 효과보다는 합의값을 계산하는데 필요한 기관별 변동성을 어떻게 보정할 것인가가 주요 관심사였다. Scott 등 (2010)는, 이후 5번째 실험인 VIRI에서 얻은 결과로부터 Z-score라는 개념을 도입해 각 기관들의 퍼포먼스를 평가하여 기관 간의 측정 차이를 이해하려는 시도를 했다 (Scott 등, 2010). 이는 이상점을 가려내서 합의값을 계산하기 위한 절차 중 일부로 시행되었고, 이 또한 각 기관 간에 발생하는 차이를 궁극적으로 이해하려는 연구는 아니었다.

본 연구에서는 고고학적 자료에 의해 연대가 대략적으로 알려진 동일 시료들을 블라인드 테스트를 통해 여러 기관에서 여러 번 측정하여 그 결과를 비교하는 실험을 통해서, 방사성탄소연대측정에 내재되어 있는 변동성을 파악하여 수치화하고 이러한 변동성이 어디서부터 유발되는 것인지를 이해하려고 한다. 연구를 통해 변동성을 유발하는 요인과 그 크기에 대해 정확히 파악할 수 있다면 방사성탄소연대측정의 오차를 줄일 수 있는 실험을 디자인할 수 있을 것이다. 또한 본 연구에서는, 실험으로 얻어진 측정치를 바탕으로 신뢰성 있는 합의값을 계산하여, 측정된 방사성탄소연대를 이용하는 사용자 입장에서 실험을 통해 얻은 여러 개의 탄소연대측정치 중에 어떤 측정치를 이용하는 것이 적절한가에 대한 기준을 마련하고자 한다. 이를 통해 고고학계에서 방사성탄소연대측정법을 이용하는데 있어 논란이 되는 문제들을 어느 정도 해결할 수 있을 것이라 기대하며, 방사성탄소연대측정법에 대한 신뢰성을 구축하는 것에도 일조할 수 있을 것이라 생각한다.

본 논문은 다음과 같이 구성되어 있다. 제 2절에서는 방사성탄소연대측정법에 대해 간략하게 설명한다. 제 3절에서는 본 연구의 목적을 달성하기 위해 계획한 실험의 내용과 실험으로부터 얻은 자료에 대해서 설명한다. 제 4절에서는 베이지안 통계모형을 이용한 분석방법에 대해서 설명하고 제 5절에서는 결과에 대해 서술한다. 마지막으로 제 6절에서는 앞으로 남아있는 연구과제에 대해 언급하고 결론을 내어 논문

을 마무리 짓도록 한다.

2. 방사성탄소연대측정법

방사성탄소연대측정법은 대기 중의 탄소 동위원소, C12, C13, C14의 비율은 항상 일정하며 대기를 통해 호흡하는 유기체의 체내에도 이 비율은 유지된다는 사실을 바탕으로 한다. 유기체가 사망하여 호흡을 멈추게 되면서 불안정상태의 동위원소인 C14가 일정한 반감기로 감소하는데 이 원리를 이용하여, 주어진 샘플(유기물)에 잔존하고 있는 C14의 양과 다른 안정상태 원소의 양을 측정 및 비교함으로써 연대를 측정한다. 샘플의 C14 양을 측정하는 방법으로는 크게 3가지가 있다. 가스 비례 계수기(Gas Proportional Counting; GPC), 액체 섬광 계수기(Liquid Scintillation Counting; LSC) 그리고 가속기 질량 분석(Accelerator Mass Spectrometry; AMS)이 그것인데, 이 중 가장 현대적인 방법이며 정확하게 탄소의 양을 측정할 수 있는 방법은 가속기 질량 분석 방법이다. 가속기 질량분석 방법은 시료 안의 탄소원자를 직접 세서 동위원소의 비율을 계산함으로써 C14의 양을 측정하는 방법으로, 시료 속에 포함된 C14가 자연적으로 붕괴하면서 방출하는 전자의 개수를 세서 탄소의 양을 알아내는 고전적인 방법보다 훨씬 정확한 측정이 가능하다는 장점이 있다.

일반적으로 방사성활동(Radioactivity)이라는 용어는 동위원소 탄소의 비율(C14/C12)을 나타내는데 사용한다. 방사성탄소연대는 방사성동위원소의 반감기를 이용하여 다음과 같은 수식으로 계산할 수 있다.

$$t = -\frac{1}{\lambda} \ln \frac{A_t}{A_0}.$$

위에서 $1/\lambda$ 는 C14의 평균 생존시간을 나타내며, A_0 는 $t = 0$ 일 때의 샘플의 방사성활동(즉, 샘플이 살아있는 환경에서의 방사성활동), A_t 는 샘플의 현재 방사성활동을 나타낸다.

그러나 실제로는 A_t 와 A_0 를 정확히 알 수 없기 때문에 실험실에서 측정된 값들을 이용하여 다음의 식으로 탄소연대를 계산하게 된다 (Stuiver와 Polach, 1977).

$$t = -8033 \ln \frac{A_{SN}}{A_{ON}}. \quad (2.1)$$

위에서 A_{SN} 은 샘플로부터 측정된 방사성활동을 1950년을 기준으로 보정한 값을 뜻하며, A_{ON} 은 현대기준(Modern standard)의 방사성활동을 1950년 기준으로 보정한 방사성활동을 의미한다. 이를 보정하는데 배경샘플(Background sample)에서 측정된 방사성활동이 이용된다. 이렇게 계산된 t 를 일반 방사성탄소연대(Conventional Radiocarbon Ages; CRA)라 부르며 BP(Before Present)로 표기한다. 2차대전시기에 핵실험에 의해 대기 중의 C14양이 인위적으로 변화한 시점이 약 1950년으로 알려져 있으며 대부분의 연대측정시료들은 이 이전의 연대를 갖기 때문에, 1950년을 기준으로 삼으며 이 해를 0 BP로 나타낸다.

현대기준으로는 옥살산 1(Oxalic Acid 1)이라는 국제적인 표준샘플이 주로 이용되는데, 이는 1950년, 즉 0BP에서의 방사성활동의 기준이 되는 샘플을 뜻한다. 배경샘플은, 방사성탄소의 전문가들이 ‘활동이 없다(no activity)’ 또는 ‘남아있는 C14의 양이 0이다’라고 확신하는 샘플을 뜻한다. 이들은 일반적으로 5만년 이상으로 오래된 지질학적 샘플로 구성된다. 가속기 질량분석 방법으로 탄소연대를 측정하기 위해서는 먼저 물리적, 화학적 전처리를 통해 시료를 오염시킬 수 있는 물질을 제거해야한다. 이러한 전처리 과정을 거친 배경샘플의 방사성활동을 측정하게 되면, 전처리 과정에 의해서 오염되는 부분으로 인해 과대 계산되는 방사성활동을 측정할 수 있다. 이를 샘플과 현대기준 샘플의 방사성활동에서 공제함으로써 전처리과정으로 인해 오염되지 않은 샘플의 방사성활동을 계산한다.

식 (2.1)로부터 계산된 BP는 A_{SN} 과 A_{ON} 을 측정하는데서 발생할 수 있는 오차 때문에 변동성을 갖게 된다. 일반적으로 각 탄소연대측정기관에서 BP로 보고되는 표준편차(Standard Deviation; SD)는 위 두 측정치로부터 유도되는 변동성이 된다.

BP를 보고할 때에는 몇 가지 관습적인 용례가 사용된다. 만약 BP가 200년 이내의 것이면, ‘현대(Modern)’라는 용어를 사용한다. 만약 샘플의 나이가 1950년 이후의 것이면, ‘>현대’라는 기호를 사용한다. 현대와 배경샘플 사이의 연대를 갖는 샘플의 연대는 BP를 써서 표기하면 되는데, 이들은 BP가 속하는 연대의 범위에 따라 일정한 규칙으로 반올림하여 이용해도 무방하다.

3. 실험계획 및 자료

샘플의 정확한 실제연대를 알 수 없기 때문에 방사성탄소연대측정법을 통해서 계산된 탄소연대측정치 는 실제연대를 중심으로 하는 어떠한 분포로부터 생성된 자료로 간주하게 된다. 국제방사성탄소연대비 교 실험에서 얻은 자료들의 상자그림을 보면 몇몇 이상점들이 보이기는 하나 정규분포 가정을 하는 것에 큰 무리는 없는 것으로 보이므로, 본 연구에서 탄소연대측정값을 실제연대를 중심으로 하는 정규분포를 따르는 확률변수라고 가정하는 것이 적절하다고 판단할 수 있다 (Rozanski 등, 1992; Scott 등, 2010; Scott, 2003). 즉,

$$X \sim N(\mu, \sigma^2)$$

으로 가정한다. 여기서 X 는 샘플의 탄소연대측정값을 나타내는 확률변수이며 μ 는 샘플의 실제연대, 그리고 σ^2 은 탄소연대측정법에 내재되어 있는 불확실성, 또는 탄소연대측정치의 오차를 설명할 수 있는 변동성으로 간주할 수 있다.

본 연구에서 우리의 주요 관심사는, 탄소연대측정법에 내재되어 있는 불확실성을 야기하는 요인과 이 변동성을 측정하는 것이다. 변동성을 야기하는 요인 중에는, 측정기관 간에 존재하는 측정오차, 측정기관 내에서 존재하는 측정오차, 동일 유적에서 얻어진 샘플들의 종류에 따른 오차 등이 있을 수 있다. 이러한 요인들에 대한 이해를 통해서 변동성을 모형화하고 고고학적 연대가 추정된 탄소연대와 얼마나 일치 하는가에 대해 판단할 수 있다.

블라인드 테스트에 이용된 시료는 두 곳의 고고학 유적으로부터 채집되었다. 한 곳은 출토 토기에 의할 경우 서기 3세기 중반대로 추정되는 연천 남계리 유적(서울대학교박물관)으로서, 하나의 주거지에서 탄 화목재 3점이 채집되었다. 또 한 곳의 유적은 서울 광진구 아차산에 위치한 고구려 요새유적인 흥련봉 2보루(고려대학교 고고환경연구소)로서, 이곳에서는 탄화목재 2점이 대상으로 선정되었는데, 발굴단은 고구려 한강유역 진출 문헌기사를 고려하여 이 유적이 대체로 6세기 전반 경에 해당할 것으로 추정하고 있다. 이와 같이, 2개소에서 총 5점의 탄화목재가 채집되어 실험에 이용되었는데, 이 시료들은 수중분 석 결과 모두 참나무속에 속하는 것으로 판정되었다. 용어상의 혼란을 피하기 위해 여기서는 이 5점의 샘플을 메가샘플이라 지칭하기로 한다.

실험은 5개의 메가샘플을 각각 16개씩으로 나누어 총 80점의 샘플을 5개의 AMS기관들(서울대 기초과 학공동기지원, 미국 Beta Analytic Inc., 일본 Paleo Labo, 미국 Illinois State Geological Survey at University of Illinois, 영국 Radiocarbon Dating Facility at Queens University, Belfast)에 시차를 두고 반복적으로 보내는 방식으로 진행되었다. 블라인드 테스트의 효과를 높이기 위해, 각 측정기관에는 샘플의 고고학적 정보를 최소한만 제공하였다. 동일 유적에서 얻은 샘플들은, 각기 다른 AMS기관에서 그 연대를 측정하더라도 거의 비슷한 연대가 측정될 것으로 여겨진다. 이러한 가정에도 불구하고, 만약 실제 자료에서 측정된 탄소연대측정치가 차이를 보인다면, 탄소연대측정치를 변화시키는 어떠한 요인이

Table 3.1. Structure of data

	서울대	Beta	Paleo Labo	Illinois	Queen's Belfast	합계
남계리1	4	4	4	4	4	20
남계리2	1	2	2	2	2	9
남계리3	2	2	2	2	2	10
홍련봉1	4	4	4	4	4	20
홍련봉2	4	4	4	4	4	20
합계	15	16	16	16	16	79

있는 것으로 판단하고 이를 모형화하여 탄소연대측정에 존재하는 불확실성에 대한 정보를 얻을 수 있을 것이다.

80개의 샘플을 AMS기관에 보내 총 79개의 탄소연대측정치를 얻었다. Table 3.1은 자료의 전체적인 구조를 보여준다. 연천 남계리의 메가샘플 1점과 홍련봉의 메가샘플 2점에서는 5개의 AMS기관에서 각 4번씩 반복 측정된 탄소연대측정치를 얻었으며, 나머지 2점의 메가샘플에 대해서는 기관마다 각 2번씩 반복 측정된 탄소연대측정치를 얻었다. 이 중 측정불가로 보고된 결측치 1개를 제외하면, 모형에 이용할 수 있는 자료는 총 79개이다.

기관으로부터 보고된 자료에서 우리가 관측치로 이용한 값은 BP이며, 이를 탄소연대측정치로 생각한다. 하지만 고고학적으로 BP는 그 자체로서의 의미는 적으며, 일반적으로는 고고학계에서 이용하는 보정곡선(Calibration curve)이라는 추정곡선을 이용하여 BP로부터 환산한 cal BP(Calendar BP)를 실제 연대추정치로 이용한다. 그럼에도 불구하고 본 논문에서 BP를 관측치로 다룬 이유는, 사용자의 입장에서 여러 개의 측정치를 어떻게 혼합해서 이용할 것인가에 대한 문제와 관련이 있다. 동일한 샘플에 대해 여러 AMS기관에서 보고된 cal BP를 이용하는 경우에는, 하나의 합의값을 도출하는데 있어 보정곡선 자체에 존재하는 불확실성을 다시 고려해야 한다는 문제가 발생하게 된다. 만약 BP를 이용해 먼저 하나의 합의값을 계산하게 된다면, 대응되는 하나의 cal BP를 도출할 수 있게 된다. 이는 하나의 합의값을 도출하는데 있어 발생할 수 있는 불확실성은 고려하지 않아도 된다는 것을 의미하며, 따라서 본 논문에서는 BP를 이용하여 합의값을 계산하는 것을 보다 합리적이라고 여겨 관측치로 사용하였다.

수집된 자료의 개괄적인 형태는 Figure 3.1과 같다. 이를 통해 어떤 유적에서 채집한 메가샘플인지에 따라 측정된 BP값에 차이를 보이는 것을 알 수 있다. 그리고 동일한 메가샘플에 대해서도 서로 다른 AMS기관에서 측정한 BP값에 차이가 있다는 것을 확인할 수 있다. 또한, 각 기관에서 반복측정한 자료로부터 기관마다 측정된 BP값이 갖는 범위가 다를 수 있다. 79개의 자료 중에는 이상점으로 판단되는 2개의 측정치가 존재했으나 일단은 이 측정치를 모두 적절한 자료로 간주하여 이용하였다.

4. 통계모형

수집한 자료의 특징을 파악하여, 탄소연대측정치에 대해 다음과 같은 통계모형을 가정했다. 탄소연대측정치의 모평균이 어떤 연대의 유적에서 얻은 자료인가, 어느 측정 기관에서 측정하는가, 어떤 메가샘플을 선택하였는가에 따라 변화할 수 있다고 생각되므로 이를 모형에 반영하였다.

$$\begin{aligned}
 y_{1jkl} &= \mu_1 + \alpha_j + \beta_k + \epsilon_{1jkl}, & j = 1, \dots, 5, k = 1, 2, 3, \\
 y_{2jkl} &= \mu_2 + \alpha_j + \beta_k + \epsilon_{2jkl}, & j = 1, \dots, 5, k = 4, 5, \\
 \epsilon_{1jkl}, \epsilon_{2jkl} &\sim N(0, \sigma_j^2).
 \end{aligned}$$

위 모형에서 y_{1jkl} 과 y_{2jkl} 은, 각각 남계리와 홍련봉의 샘플로부터 j 번째 AMS기관과 k 번째 메가샘플에

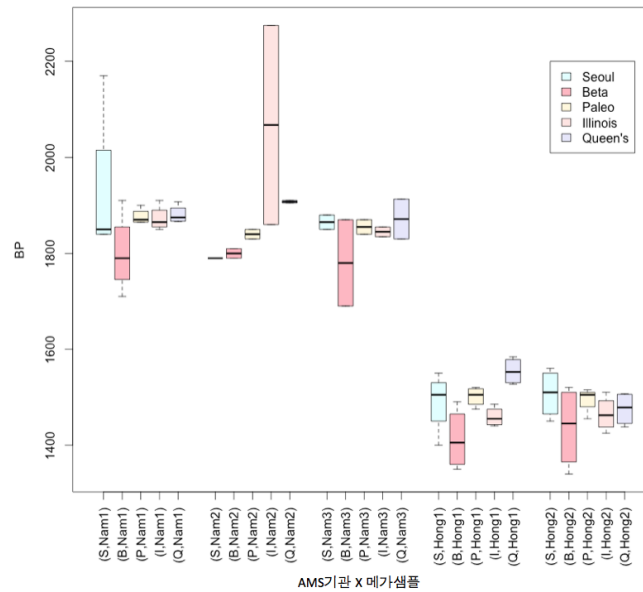


Figure 3.1. Boxplot of data.

서 측정된 탄소연대측정치를 의미한다. l 은 반복측정을 나타낸다. μ_1 과 μ_2 는 연천군과 홍련봉, 각각의 유적에서의 탄소연대측정치 평균을 뜻하며 이는 모수효과(Fixed effect)로 가정한다. α_j 는 AMS기관이 탄소연대측정치에 미치는 영향, β_k 는 각 유적에서의 메가샘플의 종류가 탄소연대측정치에 미치는 영향을 의미하며 변량효과(Random effect)들로 생각한다. ϵ 은 반복적으로 측정된 탄소연대측정치에 차이를 유발하는 오차를 의미한다. 위 모형에서 ϵ 은 서로 독립이며 측정기관에 따라 차이가 있다고 가정하였다. 변량효과인 α_j, β_k 에는 다음과 같은 가정이 주어진다.

$$\alpha_j \sim N(0, \sigma_A^2), \quad j = 1, 2, \dots, 5,$$

$$\beta_k \sim N(0, \sigma_B^2), \quad k = 1, 2, \dots, 5,$$

σ_A, σ_B 는 각각 측정기관 간의 변동성과 메가샘플 간의 변동성을 의미한다.

관측치들은 동일한 유적으로부터 얻은 샘플에 대해 서로 다른 기관에서 측정된 자료로 생각할 수 있다. 또는, 동일한 기관에서 측정된 서로 다른 샘플의 측정치로 볼 수도 있다. 따라서 자료에 대한 모형을 동일한 모수들로부터 파생되는 계층적인 구조로 나타내는 것이 적절할 것이다. 또한 본 연구에서 이용된 자료의 경우, 고고학적 자료에 의해 샘플들의 연대를 어느 정도 알고 있기 때문에 샘플의 연대에 대한 사전정보가 있다고 간주하고 주관적 사전분포(Subjective prior)를 사용할 수 있다. 따라서 계층적 모형의 추정에 유용성을 가지며 사전정보를 모형에 이용할 수 있는 베이저안 방법론을 이용하는 것이 적절하다고 판단하였다.

위 모형에서 추정해야 할 모수는 μ, α, β , 그리고 $\sigma, \sigma_A, \sigma_B$ 이다. 전문가가 판단하는 유적의 대략적인 연대로 μ_1 과 μ_2 의 사전분포를 구성하였다. 남계리 유적과 홍련봉 유적의 대략적인 연대를 BP로 환산하면, 95% 사전분포의 구간을 각각 2000~1600 BP, 1500~1400 BP로 생각할 수 있다. 이를 이용해 정규분포 가정을 한 사전분포의 평균과 표준편차를 구해보면 μ_1 의 경우 1800, 100, 그리고 μ_2 의 경우 1450, 25이 된다. 전문가들이 자신의 지식을 과신하는 경향이 있으므로 실제 사용된 사전분포의 표준편

Table 5.1. Estimated mean, median, standard deviation and 95% confidence interval of each parameter

	Mean	Median	SD	95% credible set
μ_1	1859.41	1859.60	14.35	(1830.48, 1886.62)
μ_2	1492.03	1492.26	15.03	(1460.91, 1521.27)
α_1	6.80	5.72	14.85	(-20.38, 39.60)
α_2	-29.73	-27.36	19.70	(-74.02, 1.60)
α_3	3.16	2.47	11.80	(-18.85, 28.94)
α_4	3.84	3.04	15.50	(-26.02, 38.25)
α_5	17.64	16.33	13.58	(-5.89, 47.97)
β_1	7.18	6.64	9.11	(-8.95, 26.81)
β_2	-2.90	-2.67	9.53	(-22.73, 15.27)
β_3	-4.35	-4.25	9.47	(-22.76, 13.84)
β_4	6.04	5.65	9.98	(-12.77, 26.19)
β_5	-5.56	-5.20	10.16	(-27.51, 13.42)
σ_1	83.64	81.71	14.68	(60.77, 117.23)
σ_2	74.94	72.08	15.89	(51.45, 112.14)
σ_3	20.17	19.82	3.80	(14.36, 28.97)
σ_4	102.59	100.55	17.66	(74.47, 142.67)
σ_5	36.52	35.48	6.85	(25.68, 52.73)
σ_A	18.40	16.41	8.79	(7.10, 40.28)
σ_B	11.15	10.37	4.01	(6.03, 21.16)

차는 100, 25의 2배로 하였다. 따라서 모형에 사용된 사전분포는

$$\mu_1 \sim N(1800, 200^2), \quad \mu_2 \sim N(1450, 50^2)$$

와 같다. 계산을 쉽게 하기 위해, $\sigma, \sigma_A, \sigma_B$ 에 대해서는 정규분포와 쥘레(Conjugate)인 역감마(Inverse Gamma)분포를 가정하였다. 본 논문에서는 탄소연대측정의 변동성을 파악하기 위해서 역사적 자료에 의해 유적의 대략적인 실제 연대를 알고 있는 자료를 이용하였으며, 그러므로 샘플에서 탄소연대측정치 오차가 최대 200년을 넘지 않을 것이라는 사전정보를 가지고 있다. 이러한 사전정보를 바탕으로 각 분산에 대한 사전분포를 설정하였으며, 따라서 역감마 분포의 초모수(Hyperparameter)는 사전분포의 민감성 테스트를 통해 다음과 같이 설정하였다.

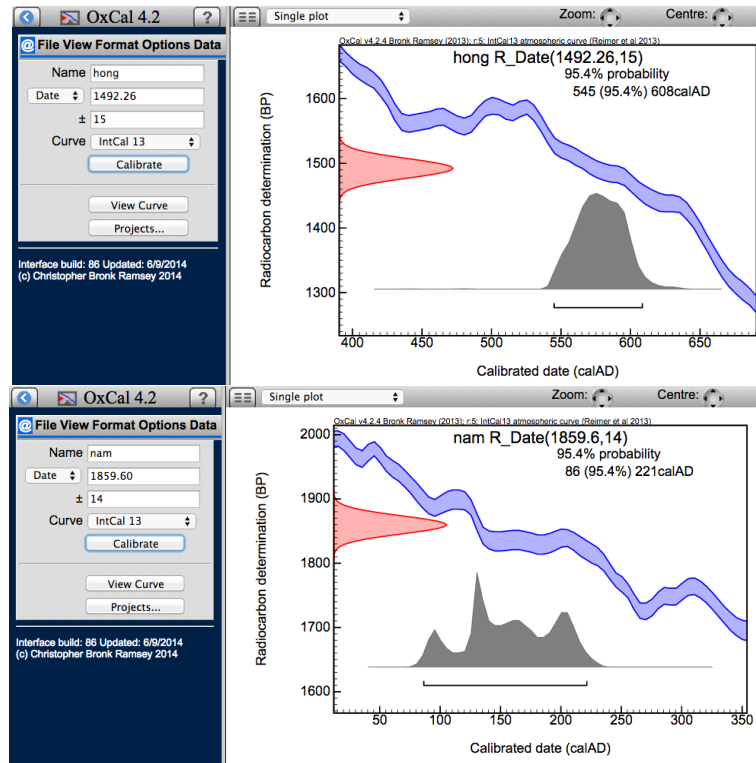
$$\sigma^2, \sigma_A^2, \sigma_B^2 \sim IG(2, 200).$$

5. 통계분석의 결과

베이저안 방법론에서는, 알고자 하는 모수에 대한 사전정보가 자료에 의해 갱신되는 경우에 모수에 대한 정보가 어떻게 변하는가에 관심을 갖는다. 모수에 대한 사전정보를 분포로 나타낸 것을 사전분포(Prior distribution), 그리고 업데이트 된 정보를 분포로 나타낸 것을 사후분포(Posterior distribution)라고 부른다. 베이저안 방법론을 이용한 모수의 추정, 사후분포를 통해 이루어진다. 일반적으로 사후분포의 정확한 형태를 알기 어려우므로, 마코프체인 몬테카를로(Markov Chain Monte Carlo; MCMC) 방법을 이용하여 사후분포를 구성하는 샘플들을 얻고 그 샘플의 히스토그램으로 사후분포의 형태를 파악하게 된다. 즉, 마코프체인 몬테카를로 방법을 이용하면, 모수의 사후분포로부터 생성되었다고 생각할 수 있는 마코프체인 몬테카를로 샘플을 얻게 되고 이를 이용하여 모수를 추정할 수 있는 것이다. 모수의 추정치는 마코프체인 몬테카를로 샘플의 평균(mean), 최빈값(mode), 중앙값(median) 등을 이용할 수 있

Table 5.2. Estimated BP of Namkye-Ri and Hongryun-Bong sites

남계리	홍련봉
1859.60	1492.26

**Figure 5.1.** Cal BPs calculated by the consensus value of Hongryun-Bong site(top) and Namkye-Ri site(bottom).

다. 본 연구에서 얻어진 사후분포는 대칭에 가까우므로 평균, 최빈값, 중앙값 중 어느 것을 선택해도 큰 차이가 있지는 않으나, 논문에서는 자료의 극단치에 영향을 덜 받을 수 있는 중앙값을 추정치로 이용하였다.

분석은 R(ver. 3.0.3)을 이용하였으며, R에서 베이지안 모형을 분석하기 위한 프로그램인 JAGS를 이용하기 위해 ‘R2jags’ 패키지를 사용하였다. 이를 이용하면 베이지안 모형에서의 모수에 대한 마코프체인 몬테카를로 샘플을 얻을 수 있다. 최종적으로 얻어진 10,000개의 마코프체인 몬테카를로 샘플을 통해 각 모수의 사후분포의 형태를 나타낼 수 있으며, 이를 통해 추정치들을 계산할 수 있다. 전체 모수의 추정치는 Table 5.1에 나타나 있다.

5.1. 연천 남계리와 홍련봉 2보루 유적의 추정연대

추정을 통해 얻어진 연천 남계리와 홍련봉 유적에서의 탄소연대추정치는 약 1860과 1492 BP로 나타난다 (Table 5.2). 이때 유적의 실제 연도는, 보정곡선을 이용해 추정된 BP를 cal BP로 환산하여 계산할 수 있다. cal BP의 환산은, Christopher Bronk Ramsey에 의해 만들어진 OxCal 프로그램을 이용한다.

Table 5.3. Estimated BPs between AMS labs

	서울대	Beta	Paleo Labo	Illinois	Queen's Belfast
남계리	1866.21	1829.68	1862.57	1863.25	1877.05
홍련봉	1498.83	1462.30	1495.19	1495.87	1509.67

Table 5.4. Estimated BPs between megasamples

남계리1	남계리2	남계리3	홍련봉1	홍련봉2
1866.59	1856.51	1855.06	1504.87	1493.27

Table 5.5. Estimated variabilities of BP between AMS labs and between megasamples

기관	메가샘플
18.40	11.15

다(<https://c14.arch.ox.ac.uk>).

베이즈 정리에 의해서, 정규분포의 형태를 갖는 μ_1 과 μ_2 의 주변사후분포(Marginal posterior distribution)를 계산할 수 있는데 이 주변사후분포에서의 표준편차 근사값으로 마코프체인 몬테카를로 샘플의 표본표준편차를 이용할 수 있다. 본 논문에서는 근사값인 14와 15를 μ_1 과 μ_2 주변사후분포의 표준편차로 간주하였으며, 이를 OxCal에서 입력할 표준편차값으로 사용하였다. 일반적으로 탄소연대측정 기관에서는 탄소연대측정치와 탄소연대 측정치에 대한 이론적 오차를 표준편차의 형태로 제공하며, OxCal을 이용하여 실제연대를 계산하기 위해서는 하나의 탄소연대 측정치와 표준편차가 필요하다. 그러나 OxCal에서는 입력하는 탄소연대 측정치가 탄소연대의 평균이 아니라, 어떠한 평균을 중심으로 한 하나의 실현값이라고 가정하고 있다. 본 논문에서는 여러 탄소연대측정치로부터 μ_1 과 μ_2 의 마코프체인 몬테카를로 샘플을 얻어 이로부터 합의값을 구했고 이것을 하나의 유적에 대한 탄소연대측정치로 간주하였다. 이 값을 어떠한 평균을 중심으로 한 하나의 실현값으로 생각하면, μ_1 과 μ_2 의 마코프체인 몬테카를로 샘플로부터 계산된 표본표준편차를 OxCal에서 이용할 오차로 생각하는 것이 적절하다고 볼 수 있다. 따라서 Oxcal에서 이용할 각 유적의 BP연대는 남계리 1859.60 ± 14 BP, 홍련봉 1492.26 ± 15 BP이며, 이를 바탕으로 각 유적의 cal BP를 추정하였다. 이 추정은 OxCal 4.2 버전에서 이루어졌고, IntCal 13 보정곡선을 이용하였다.

보정된 탄소연대는 Figure 5.1에서 확인할 수 있다. 홍련봉 유적의 경우 AD 545-608, 남계리 유적의 경우 AD 86-221로 추정된다. 실험에 이용된 두 곳의 연대를 고고학계에서 각각 6세기 전반, 3세기 중반으로 보고 있다는 사실을 고려할 때, 본 연구에서 계산한 합의값이 실제 유적의 연대와 약간의 차이가 있다는 것을 확인할 수 있다. 그러나 이러한 차이가 발생하는 것은, 다른 시기의 유구가 기존 유구에 중첩되기 때문이며 따라서 이 연대가 고고학적 연대와 비교할 때 잘못 추정된 연대가 아니라고 해석할 수 있다 (Choi, 2014).

5.2. 기관과 메가샘플 요인의 효과

위의 분석은 모든 기관에서 측정된 값이 신뢰할 만하다는 가정 하에서 이루어졌지만, 기관별로 측정결과에 차이가 있다는 점은 고려할 필요가 있다. Table 5.3의 결과는 각 유적에 대한 5곳의 AMS기관의 탄소연대추정치를 보여주는데, 대부분의 기관이 큰 차이를 보이지 않지만 예외적으로 Beta에서의 추정치가 다른 기관들에 비해 약 40년 정도 더 최근 것으로 나타나는 경향이 발견된다. 또한, Beta는 동일시료에 대한 두 번의 측정실험 간에 큰 차이를 나타내고 있는 문제를 보이고 있다.

Table 5.6. Estimated mean, median, standard deviation and 95% credible set of each parameter after removing outliers

	Mean	Median	SD	95% credible set
μ_1	1852.96	1853.39	12.48	(1826.40, 1876.70)
μ_2	1483.14	1483.54	13.42	(1456.08, 1509.47)
α_1	2.93	2.39	11.95	(-19.87, 28.71)
α_2	-27.19	-25.89	17.06	(-63.57, 1.85)
α_3	10.17	9.44	10.53	(-8.84, 33.44)
α_4	-6.34	-6.70	10.56	(-26.66, 15.32)
α_5	22.71	21.93	12.92	(-0.71, 50.98)
β_1	8.53	8.10	8.52	(-7.47, 26.54)
β_2	-5.48	-5.24	8.99	(-24.93, 11.54)
β_3	-3.00	-2.81	8.58	(-21.09, 14.03)
β_4	2.64	2.36	9.08	(-15.68, 22.15)
β_5	-3.35	-3.33	9.02	(-21.46, 14.45)
σ_1	42.74	41.77	7.90	(30.54, 60.89)
σ_2	72.51	70.38	14.34	(50.69, 105.84)
σ_3	19.80	19.28	3.58	(14.23, 28.11)
σ_4	26.84	26.26	5.04	(18.81, 38.65)
σ_5	38.18	37.21	7.13	(26.83, 54.23)
σ_A	17.96	16.67	7.61	(7.68, 36.13)
σ_B	10.83	10.06	3.88	(5.88, 20.41)

한편, Table 5.4는 5점의 메가샘플마다 추정된 탄소연대를 보여준다. 이는 메가샘플이 탄소연대측정치에 미치는 효과를 나타내는 것으로, 그 효과는 약 10년 내로 미미함을 확인할 수 있다. 따라서 메가샘플의 선택, 즉 동일한 유적 또는 유구에서 어느 시료를 고르느냐의 문제는 적어도 본 실험에 있어서는 유적의 연대를 예측하는데 있어 큰 영향을 미치지 않았다고 볼 수 있다.

기관과 메가샘플이라는 각 요인이 탄소연대측정치를 변동시키는 정도는, Table 5.5에서 확인할 수 있다. 5곳의 AMS기관인 서울대, Beta, Paleo Labo, Illinois, Queen's Belfast의 표준편차는 각각 약 84, 74, 20, 103, 36년으로 추정된다. 이는 각 기관에서 탄소연대측정을 할 때 발생할 수 있는 오차의 크기를 의미한다. 서울대와 Illinois의 경우 변동성이 매우 크다는 것을 확인할 수 있다.

5.3. 이상점의 영향

남계리에서의 두 결과는(2275 ± 25 Illinois, 2170 ± 60 서울대) 다른 측정결과와 큰 차이를 보였다. 추정의 정확도를 높이기 위해서는 이들이 실제로 이상점인지 아닌지를 판단하여 이를 제외 또는 수정하여 이용할 필요가 있다. 흥련봉에서도 극단적인 수치를 나타내는 측정치가 발견되었으나, 남계리에 비해 상대적으로 차이가 크지 않아 본 분석에서는 이들 수치를 이상점으로 판단하지 않았다.

이상점을 제외한 후 위와 동일한 방법으로 모수를 추정하였으며, 추정결과는 Table 5.6에서 확인할 수 있다. 이를 이상점을 제외하기 전의 추정결과와 비교해 볼 때 가장 눈에 띄는 점은, 서울대와 Illinois의 변동성이 크게 줄어들었다는 점이다. 이상점을 제외한 후의 각 기관별 변동성의 추정치는 각 기관에서 제공한 보고서에서 보고된 오차와 거의 유사한 값을 보이고 있다 (Table 5.8). 이 오차는 제 2절에서 언급한, 이론적인 탄소연대측정오차를 의미한다고 볼 수 있다. 따라서 이 오차를 줄이기 위해서는 식 (2.1)의 A_{SN} 과 A_{ON} 의 변동성을 줄이는 방법이 필요하다고 여겨진다. 그러나 이러한 통계적인 변동성

Table 5.7. Estimated variability of BP in each AMS lab

	서울대	Beta	Paleo Labo	Illinois	Queen's Belfast
이상점 제외 전	83.64	74.94	20.17	102.59	36.52
이상점 제외 후	42.74	72.51	19.80	26.84	38.18

Table 5.8. Range of variabilities of BP reported by each AMS lab

서울대	Beta	Paleo Labo	Illinois	Queen's Belfast
40 ~ 60	30	15 ~ 20	15 ~ 25	25 ~ 42

을 줄이는 것은 이 글의 논의범위를 벗어나므로 더 이상 다루지 않는다.

반면 BP추정치, 또는 cal BP추정치(홍련봉의 경우 124~226 AD, 남계리의 경우 550~615 AD)는, 이상점이라 여겨지는 자료를 제외하여도 크게 변하지 않는 경향이 보인다 (Table 5.6). 따라서 모수에 대한 마코프체인 몬테카를로 샘플의 중앙값을 모수의 추정량, 즉 탄소연대의 합의값으로 이용하는 것이 적절하다고 판단할 수 있다.

6. 결론

연구에서는 탄소연대측정법의 신뢰성을 파악하기 위한 실험을 시행하고, 실험결과를 바탕으로 탄소연대 측정치를 변화시킬 수 있는 요인과 그 변동성의 크기에 대해 통계학적으로 분석하였다. 또한 여러 개의 탄소연대측정치들을 이용해서 합의값을 얻는 통계적 방법에 대해서도 제안했다. 그 결과, 마코프체인 몬테카를로 샘플들의 중앙값으로 탄소연대의 합의값을 추정할 수 있었으며, AMS기관 내, 기관 간, 그리고 동일유적에서 채취한 메가샘플 간에 변동성이 존재한다는 것을 확인하고 그 수준을 나타낼 수 있었다.

본 연구는 탄소연대측정치를 이용하는 사용자 입장에서 변동성이 발생할 수 있는 요인을 파악하고 그 변동의 정도가 어느 정도인지를 실험을 통해 수치적으로 나타냈다는 데에 의미가 있다. 그러나 기관이 갖는 측정의 변동성은 최대 70 BP가 될 수 있다는 사실 또한 밝혀졌다. 이는 탄소연대측정치의 변동성이 탄소연대측정치에 비해 상대적으로 크다는 것을 의미한다. 만약 몇몇 기관에서의 실제 탄소연대측정에 내재되는 오차가 이와 같다면, 삼국시대와 같이 비교적 좁은 범위 안에서 편년이 결정되는 경우에서 탄소연대측정법을 통한 정확한 연대추정은 고고학자에게 고민으로 다가올 수밖에 없다.

그러나 여러 기관에서 반복적으로 측정한 탄소연대측정치를 이용하여 베이지안 방법으로 얻은 합의값의 경우, 변동의 수준이 약 15년 이내로 줄어든 것을 확인할 수 있었다. 이는, 여러 개의 탄소연대측정치로부터 얻은 합의값의 변동은 실제 각 기관이 가지는 변동보다 훨씬 작으며, 따라서 합의값을 이용한 탄소연대추정의 신뢰성이 어느 정도 보장됨을 뜻한다. 따라서 탄소연대측정치로부터 연대를 추정하기 위해서는 하나의 탄소연대측정치만이 아니라 여러 개의 측정치로부터 얻은 합의값을 이용하는 것이 바람직하다.

또한 본 연구 결과에서는, 이상점을 제외하였을 때의 기관 내 변동성이 실제 각 기관에서 보고하는 변동성과 거의 유사하다는 결론을 얻을 수 있었으며, 따라서 기관마다 보고되는 탄소연대측정의 변동성은 신뢰할 만하다는 사실을 알 수 있었다.

탄소연대측정치를 이용하는 사용자 입장에서는, 오차를 실제로 고고학의 분야에서 받아들일만한 수준으로 줄이기 위해서 어떠한 기관을 선택하여 어떤 실험 디자인 하에서 방사성탄소연대측정치를 얻어야 하는지에 관심이 있을 것이다. 연구의 결과를 보면, 보고되는 탄소연대측정의 변동성이 매우 작은 기관 하

나를 선택하여 여러 개의 탄소연대측정치들을 얻고 합의값을 통해 탄소연대를 추정하는 것이, 여러 기관으로부터 여러 탄소연대측정치들을 얻어 탄소연대를 추정하는 것보다 효율적이며 정확하기까지 하다는 생각을 들게 만들지도 모른다. 그러나 이러한 실험계획은 주의를 기울여야 한다. 만약 특정 기관의 탄소연대측정이 기계적 결함 등으로 인해서 편향성을 가지는 경우에는, 비록 그 오차가 작다고 하더라도 추정된 탄소연대가 실제 연대와는 매우 다를 수 있기 때문이다.

그러나 본 연구만으로는 기관의 편향성을 판단하기 어려운 것이 사실이다. 본 연구가 연천 남계리와 흥련봉 유적의 고고학적 연대를 확립하는 과학적 근거가 될 수 있다면, 본 자료를 정확한 실제연대를 알고 있는 경우의 자료로 간주하여 몇몇 기관들의 편향성 판단에 대한 진전연구가 이루어 질 수 있을 것이며, 이를 통해 연대추정에서 신뢰성 있는 합의값을 얻는 실험설계를 제안할 수 있다. 앞으로의 연구를 통해, 한국 고고학계에서 탄소연대측정법을 이용한 연대추정의 신뢰성을 본격적으로 검증하고, 적절한 실험설계를 통해 결과를 올바르게 해석할 수 있는 기틀이 마련되리라 기대한다.

References

- Choi, J. (2014). Interpretation of archaeological data in Korean historical archaeology and radiocarbon dating, *A Sourcebook for the 38th Annual Meeting of the Korean Archaeological Society*, 364–371
- Kim, J. (2014). Chronology and understandings of formal variability in Korean archaeology, *Journal of Korean Ancient Historical Society*, **83**, 5–32.
- Rozanski, K., Stichler, W., Gonfiantini, R., Scott, E. M., Beukens, R. P., Kromer, B. and VanderPlicht, J. (1992). The IAEA C14 inter comparison Exercise 1990, *Radiocarbon*, **34**, 506–519.
- Scott, E. M., Cook, G. and Naysmith, P. (2010). The fifth international radiocarbon intercomparison (VIRI): An assessment of laboratory performance in stage 3, *Radiocarbon*, **52**, 859–865.
- Scott, E. M. (2003). The third international radiocarbon intercomparison (TIRI) and the fourth international radiocarbon intercomparison (FIRI), *Radiocarbon*, **45**, 135–328.
- Stuiver, M. and Polach, H. A. (1977). Discussion: Reporting of ^{14}C data, *Radiocarbon*, **19**, 355–363.

방사성탄소연대측정법의 불확실성에 대한 베이지안 분석

이영선^{a,1} · 이재용^a · 김장석^b

^a서울대학교 통계학과, ^b서울대학교 고고미술사학과

(2014년 12월 15일 접수, 2015년 2월 16일 수정, 2015년 2월 16일 채택)

요약

본 연구는, 방사성탄소연대측정방법에 내재되어 있는 변동성을 유발하는 요인과 그 크기를 파악하고, 측정치들을 기반으로 한 고고학적 연대의 합의값을 도출하는 것에 그 목적이 있다. 이를 위해 동일한 샘플을 서로 다른 기관에 보내어 얻는 실험을 시행하였고, 실험을 통해 얻은 자료는 베이지안 방법을 이용해서 분석하였다. 분석결과로부터 측정기관, 메가샘플의 종류가 방사성탄소연대측정방법의 변동성을 야기하는 요인이 될 수 있음을 판단할 수 있었으며, 각 요인으로 인한 변동성의 크기는 약 10년 내외임을 확인했다. 기관 내에 존재하는 탄소연대측정의 변동성은 각 기관에서 보고하는 측정오차로 간주해도 무방하다는 사실 또한 밝혀냈다. 자료의 실제 연대를 알고 있다는 가정 하에서, 여러 개의 탄소연대측정치로부터 도출된 합의값은 실제 샘플의 연대와 일치했으며, 합의값의 변동성은 개별적인 탄소연대측정치의 변동성보다 훨씬 줄어드는 현상을 관찰할 수 있었다.

주요용어: 방사성탄소연대측정법, 고고학적 연대, 베이지안 통계, 합의값, 변동성

이 논문은 2013년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2013S1A5B6 043901).

¹교신저자: (151-747) 서울시 관악구 관악로 1, 서울대 통계학과. E-mail: lysstat@gmail.com