



한국심리학회지: 일반

Korean Journal of Psychology: General

2021, Vol. 40, No. 4, 567-596

<http://dx.doi.org/10.22257/kjp.2021.12.40.4.567>

베이지안 분석에서 사전분포의 이해와 적용*

이 지 윤

김 수 영[†]

이화여자대학교 심리학과

최근 베이지안 추정 방법이 사회과학 분야에서 많은 관심을 받고 있다. 베이지안 방법에는 연구자의 배경지식을 추정에 반영할 수 있는 사전분포라는 특별한 요소가 있으며, 이를 어떻게 지정하는지가 추정 전반에 영향을 미친다. 사전분포는 베이지안 분석에서 가장 중요한 요소임에도 불구하고, 사전분포를 이해하고 적절히 지정하기 위해 참고할 수 있는 방법론적 연구는 부족한 상황이다. 본 연구는 사전분포의 중요성과 사전분포 지정에 대한 전반적인 내용을 다룬다. 먼저, 연구자가 사전분포를 직접 지정하지 않는, 즉 프로그램이 제공하는 디폴트 사전분포 방법을 알아본다. 자주 사용되는 프로그램들의 디폴트 사전분포를 알아봄과 더불어 디폴트 사전분포의 알려진 문제점도 확인한다. 다음으로는 연구자가 사전분포를 직접 지정하는 방법에 대해 다룬다. 직접 지정할 수 있는 사전분포에는 무정보 사전분포와 정보 사전분포가 있으며, 어떤 사전분포를 이용할지는 모수에 대한 사전 정보의 유무에 따라 결정된다. 무정보 사전분포의 필요성과 이를 지정할 때 참고할 수 있도록 제안된 방법을 다루고, 정보 사전분포를 지정할 때 참고할 수 있는 연구들을 제공하며, 여러 연구의 기준을 종합해 연구자의 정보성 선택에 참고할 수 있는 기준을 탐색한다. 이후 본문에서 논의한 방법들을 적용한 자료 예시를 통해 실질적 도움을 제공하고자 하였으며, 마지막으로 본 연구의 의의와 한계에 대해 논의한다.

주요어 : 베이지안 방법, 사전분포, 디폴트 사전분포, 무정보 사전분포, 정보 사전분포

* 이 논문은 2019년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2019S1A5A2A03041362).

[†] 교신저자: 김수영, 이화여자대학교 심리학과, 서울시 서대문구 이화여대길 52

Tel: 02-3277-3792, E-mail: suyong.kim@ewha.ac.kr

전통적으로 구조방정식모형(structural equation modeling, SEM)의 추정에는 최대우도(maximum likelihood, ML) 방법이 주로 사용되어 왔으나(Hox & Bechger, 1998; Hoyle, 2000), 최근에는 베이지안(Bayesian) 방법이 많은 관심을 받고 있다(van de Schoot et al., 2015). 베이지안 방법은 최대우도 방법과 달리 모수에 대한 연구자의 배경지식을 추정의 과정에 반영할 수 있다(Muthén & Asparouhov, 2012). 배경지식은 사전 정보, 믿음(belief) 등으로도 불리며, 사전분포(prior distribution)라는 요소를 통해 추정에 포함된다. 연구자는 모수에 대한 정보와 그 정보에 대한 확신의 정도(degree of certainty)를 바탕으로 사전분포를 설정하는데, 정확한 사전분포는 표본크기가 작은 연구에서도 추정의 정확성을 높일 수 있는 반면(McNeish, 2016; Miočević et al., 2021), 부정확한 사전분포는 잘못된 추정치를 도출할 수 있으므로 사전분포의 설정은 항상 신중해야 한다(Depaoli, 2014; Muthén & Asparouhov, 2012). 그러나 일반적으로 베이지안 방법을 이용한 해외 및 국내 연구를 살펴보면 사전분포가 베이지안 추정에서 큰 영향을 미치는 요소임에도 불구하고, 사전분포의 개념만 설명할 뿐 실질적으로 이를 어떻게 효과적으로 지정하고 사용해야 하는지에는 초점을 맞추지 않고 있다. 베이지안 통계의 한 분야인 객관적 베이지안(objective bayesian)에서는 객관적인 사전분포(objective prior)를 제공하기 위한 연구를 진행하고 있으나, 연구자가 직접 지정하는 사전분포에 대한 연구는 거의 없다. 이로 인하여 연구자들은 프로그램에서 제공하는 디폴트 사전분포에 너무 의존하고 있으며(Smid et al., 2020a; van Erp et al., 2018), 사전 정보를 가지고 있어도 이를 사전분포에 어떻게 반영해야 할지 잘 모른다.

이에 본 연구는 이미 국내외에 많이 존재하는 베이지안 방법에 대한 개략적인 소개가 아닌, 베이지안 추정을 사용하는 데 있어서 사전분포의 지정에 초점을 두고 논의를 전개한다. 가장 먼저, 사회/과학 분야의 연구자들이 알맞은 사전분포의 지정을 위해 알아야 할 프로그램 디폴트 사전분포의 종류 및 문제점을 다루고, 다음으로는 사전분포의 핵심적인 개념인 무정보 및 정보 사전분포를 어떻게 이해하고 설정할 것인지에 대하여 논의한다. 또한 사전분포를 지정하는 예시를 보임으로써 실제 분석에서 베이지안 방법을 사용하고자 하는 연구자에게 유용한 지침도 함께 제시한다.

통계적 추론(statistical inference)에는 크게 두 가지 접근이 존재한다. 하나는 빈도주의(frequentist) 접근이고, 다른 하나는 베이지안 접근이다. 빈도주의 방법은 모수를 상수로 가정하고 점추정 또는 구간추정을 진행하는 반면, 베이지안 방법은 모수를 변수로 가정하고 모수의 분포를 추정한다. 최대우도 추정은 빈도주의 방법 중 대표적으로 사용되는 추정 방법으로서 불편향성, 효율성, 일관성 등의 측면에서 분명한 이점을 가지고 있지만, 큰 표본크기가 뒷받침되어야 한다는 한계가 존재한다. 한편, 베이지안 추정은 사전분포를 위해 필요한 정보를 찾는 데 다소 어려움을 겪을 수 있으나, ML 추정을 이용할 때 자주 발생하는 문제들을 피할 수 있다는 장점이 있다(Kruschke, 2011; van de Schoot et al., 2017). 예를 들어, 표본크기가 작을 때 ML 추정이 낮은 통계적 검정력으로 인해 받아들일 수 없는 결과를 보였다면(Zondervan-Zwijnenburg et al., 2017), 접근 이론을 따르지 않는 베이지안 추정은 적절한 사전분포의 지정을 통해 믿음만한 결과를 도출할 수 있다(Lee & Song, 2004;

Muthén & Asparouhov, 2012). 또한 ML 추정은 종종 수렴에 도달하지 못하지만(Lüdtke et al., 2011), 베이지안 추정에서는 수렴 문제가 상대적으로 잘 발생하지 않는다(Kim et al., 2013; Kim et al., 2020). 또 다른 예로, ML 추정에서는 분산 추정치가 음수가 되는 등의 Heywood case가 종종 발생하는 데 반해, 베이지안 추정은 제한된 범위를 가진 사전분포 지정을 통해 이러한 가능성을 없앨 수 있다(Can et al., 2015; McNeish, 2016; van de Schoot et al., 2014; Zondervan-Zwijenburg et al., 2017). 또한 베이지안 방법의 추정치인 사후분포는 중심을 기준으로 대칭이 아닌 형태도 가질 수 있는데, 이는 매개효과와 같이 정규성을 만족하지 못하는 모수를 추정할 때 큰 장점이 된다(Yuan & MacKinnon, 2009).

베이지안 방법이 가진 이론적 장점은 매우 분명하지만, 이러한 이점을 얻기 위해서는 적절한 사전분포의 지정이 필수적이다. 본 연구에서는 사전분포 지정을 크게 두 가지 방법으로 나누어 접근한다. 첫째는 프로그램에서 제공하는 디폴트 사전분포(default priors)를 선택하는 방법이다. 디폴트 사전분포는 이용자의 편의를 위해 제공되는 분포로서 실질적으로 많은 연구에서 사용되고 있다. 기존의 연구들에서는 대체로 관심 있는 모수(예, 회귀계수, 잠재평균 등)에는 사전분포를 설정할 때 참고할 수 있는 정보를 찾으려 하는 데 반해, 상대적으로 관심 없는 모수(예, 분산, 공분산, 잔차분산 등)에는 디폴트 사전분포를 할당하는 경향을 보여 왔다(Chiorri et al., 2014; Fang et al., 2019; van de Schoot et al., 2017; van Erp et al., 2018). 둘째는 연구자가 자신의 배경지식을 바탕으로 사전분포를 직접 지정하는 방법이다. 사전분포를 직접 지정하기 위해서는 사

전 정보와 그 정보가 얼마나 확실한지에 대한 믿음의 정도가 필요하다. 모수에 대한 정보는 보통 선행연구나 메타 분석, 리뷰 논문 등으로부터 얻을 수 있으며 전문가의 의견을 참고할 수도 있다(Muthén & Asparouhov, 2012). 일련의 과정을 통해 정보를 찾지 못했다면 연구자는 무정보 사전분포(noninformative priors)를 설정할 수 있다. 무정보 사전분포는 추정에 미치는 영향이 거의 없도록 지정한 사전분포이다. 반면, 모수에 대해 찾은 정보가 있다면 이를 바탕으로 정보 사전분포(informative priors)를 설정할 수 있다. 특히 모수에 대한 좋은 정보를 정확히 반영한 정밀한 사전분포(accurate priors)는 추정 결과의 정확도를 높일 수 있고(Zondervan-Zwijenburg et al., 2017), 작은 표본크기 연구의 한계를 극복할 수 있다.

일반적으로 사전분포를 설정할 때는 모수로 예상되는 값을 사전분포의 평균(또는 최빈값, 중앙값)으로 설정하고, 정보가 얼마나 확실한지를 고려하여 사전분포의 분산을 결정한다. 사전분포의 분산은 사전분포의 정보성(informativeness)을 결정하게 되며, 정보성은 일반적으로 사전 정보가 확실한 정도에 따라 무정보(noninformative), 약한 정보(weakly informative), 강한 정보(strong informative)로 나뉜다(Gelman et al., 2013; Holtmann et al., 2016; McNeish, 2016). 사전분포를 정밀하게 설정하기 위해서는 사전 정보도 필요하지만, 그 정보가 얼마나 확실한지 판단하는 것도 중요하다. 하지만 안타깝게도 현재까지 적절한 사전분포의 정보성을 결정하고자 할 때 참고할만한 연구는 충분하지 않고, 정보성을 나누는 기준도 연구마다 제각각이며, 정보성 자체에 대한 연구자들의 이해도 역시 매우 낮다. van de Schoot et al. (2017)에 따르면 최근 25년

(1990~2015) 동안 수행된 경험적 베이저안 연구 중 절반 이상이 분석에 어떤 사전분포를 사용했는지와 사전분포의 정보성에 대해 밝히지 않았다. 이러한 수치는 베이저안 추정을 사용하고 있는 연구자들이 사전분포의 지정을 어려워하며, 사전분포의 중요성 또한 간과하고 있음을 암시하는 것으로 해석할 수 있다.

본 연구의 가장 중요한 목적은 베이저안 추정에서 사전분포의 역할 및 중요성을 연구자들이 충분히 이해하도록 함으로써 베이저안 추정 방법을 정확하고 효율적으로 사용하는 데 기여하는 것이다. 이에 연구자의 믿음을 반영하지 않은 디폴트 사전분포를 지정하는 경우와 모수에 대한 연구자의 사전 지식을 바탕으로 무정보 혹은 정보 사전분포를 직접 설정하는 경우를 나누어 논의한다. 즉, 본 연구를 통해 사전분포의 전반적인 부분을 다루면서 사전분포에 대한 이해를 높이고, 연구자들에게 적절한 사전분포의 지정이 주는 결과에 대한 통합적인 시각을 제공하고자 한다.

이와 같은 목적을 달성하기 위해 먼저 베이저안 추론의 핵심인 베이저 정리를 간략히 소개하고, 사전분포를 지정할 때 고려해야 하는 요소들과 함께 베이저안 방법에서 사용되는 모수의 종류 및 사전분포의 정보성을 다룬다. 이후 프로그램 디폴트 사전분포를 자세히 알아본 뒤, 고민 없이 디폴트 분포를 사용했을 때 수반될 수 있는 문제점에 대해 논의한다. 다음으로 사전분포를 설정할 때 주요 모수(예, 회귀계수, 요인부하, 절편 등)의 사전분포 정보성 지정에 참고할 수 있도록 지난 20년 동안 수행된 베이저안 연구들을 바탕으로 직접 지정된 사전분포의 정보성 경향을 정리한다. 마지막으로 실제 자료를 바탕으로 가장 기본적인 매개모형을 통하여 베이저안 추정의 분

석 예시를 제공한다.

베이저안 추론과 사전분포

Bayes 이론과 사전분포의 역할

모수를 변수로 보는 접근에서 출발한 베이저안 방법은 모수의 분포를 추정하며, 추정을 위해 베이저 정리를 이용한다. 베이저 정리란 사전분포에 우도 함수(likelihood function)를 결합함으로써 사후분포(posterior distribution)를 도출하는 베이저안 추론의 핵심 이론이다. 사전분포는 연구자가 모수에 대해 가진 사전 정보나 배경지식을 바탕으로 설정되고, 모수일 가능성이 있는 값들로 구성된다. 수집된 자료는 우도함수로 표현되며, 자료는 우도함수를 통해서만 추정에 영향을 주어야 한다(Gelman et al, 2013). 베이저 추론은 우도함수로 사전분포를 갱신(update)함으로써 사후분포를 계산한다. 즉, 사후분포는 사전분포에 포함된 모수에 대한 사전 정보와 우도함수가 가진 자료에 대한 정보가 모두 반영된 베이저안 추론의 추정치이다.

사전분포를 설정할 때 평균은 연구자가 알고 있는 정보를 바탕으로 정하고, 분산은 그 정보를 확신하는 정도에 따라 결정한다. 이때 연구자는 사전분포를 자료 수집 전에 설정해야 한다. 사전분포 지정에 자료의 정보가 반영된다면 수집된 자료의 정보가 사전분포와 우도함수를 통해 사후분포에 두 번 영향을 미치는 더블디핑(double dipping) 문제가 발생하기 때문이다. 이는 사후분포의 분산이 과소추정되는 문제를 일으킬 수 있다(van Erp et al., 2018).

베이즈 추론은 사전분포와 우도함수를 결합하는 베이즈 정리를 통해 이루어지며, 먼저 사후분포 $p(\theta|y)$ 는 식 1과 같다.

$$p(\theta|y) = \frac{p(y, \theta)}{p(y)} \quad (1)$$

위에서 θ 는 모수, y 는 자료, $p(y)$ 는 관찰된 자료의 분포이다. 조건부 확률인 사후분포를 구하기 위해서는 모수와 자료의 결합함수인 $p(y, \theta)$ 가 필요하다. 하지만 모수는 일반적으로 알 수 없는 값이므로 연구자가 $p(y, \theta)$ 를 직접 구할 수는 없다. $p(y, \theta)$ 는 우도함수의 식으로부터 유도할 수 있으며 우도함수 $p(y|\theta)$ 는 식 2와 같다.

$$p(y|\theta) = \frac{p(y, \theta)}{p(\theta)} \quad (2)$$

위에서 $p(\theta)$ 는 사전분포이다. 식 2를 식 3과 같이 푼 뒤, 이를 식 1에 대입하면 식 4와 같은 베이즈 정리를 유도할 수 있다.

$$p(y, \theta) = p(\theta)p(y|\theta) \quad (3)$$

$$p(\theta|y) = \frac{p(y, \theta)}{p(y)} = \frac{p(\theta)p(y|\theta)}{p(y)} \quad (4)$$

식 4에서 $p(y)$ 는 상수이므로 추정 과정에서 생략할 수 있고 $\left(p(y) = \int p(\theta)p(y|\theta)d\theta\right)$, 결국 사전분포와 우도함수의 곱은 식 5처럼 사후분포와 비례관계가 된다.

$$p(\theta|y) \propto p(\theta)p(y|\theta) \quad (5)$$

베이지안 추정은 이러한 비례관계를 이용하여, 즉 사전분포와 우도함수의 곱을 통해 사후분포를 추정하게 된다.

위치 모수와 분산 모수

사전분포를 정확히 이해하기 위해서는 하이퍼파라미터(초모수, hyperparameter)를 먼저 알아야 하는데, 하이퍼파라미터는 모형의 모수가 아닌 분포의 모수를 부르는 명칭이다. 예를 들어, 사전분포로 정규분포를 사용하면, 평균과 분산이 사전분포의 하이퍼파라미터가 된다. 따라서 연구자는 사전분포를 설정할 때 하이퍼파라미터를 적당한 값으로 조정하게 된다. 자주 사용되지는 않지만, 사전분포의 모수는 특정한 값이 아닌 또 다른 분포로도 설정이 가능하며, 이때는 분포의 모수를 하이퍼파라미터가 아닌 하이퍼프라이어(초사전분포, hyperprior)라 일컫는다.

위치 모수(location parameter)는 모든 실수를 추정치로 가질 수 있는 모형의 모수들을 통칭하는 용어로, 구조 모수(structural parameter)로 불리기도 한다. 구조방정식모형의 맥락에서 대표적인 위치 모수로는 요인부하, 절편, 기울기, 잠재평균, 회귀계수 등이 있고, 위치 모수의 사전분포로는 주로 정규분포가 사용된다(van Erp et al., 2018). 정규분포는 하이퍼파라미터만으로도 분포의 모양을 파악하기 쉽고, 분포를 원하는 모양으로 설정하기도 쉬우며, 공액(결레, conjugacy)라는 특성을 갖기 때문이다. 공액란 사전분포와 사후분포가 동일한 분포를 따르게 하는 성질이며(Kaplan & Depaoli, 2013), 사후분포를 닫힌 형태(closed form)¹⁾로

1) 방정식의 해가 유한한 경우로서 근사 추정치가

계산해 베이저안 분석의 복잡한 수학적 계산 절차를 간단하게 해준다(Baldwin & Fellingham, 2013; Gelman et al., 2013; van de Schoot et al., 2017). 공액을 갖는 사전분포는 공액 사전분포(conjugate priors)라 부르며, 특별한 이유가 없는 이상 연구자는 모수에 공액 사전분포를 할당하는 것이 일반적이고, 특히 *Mplus*에서는 공액 사전분포만 설정이 가능하다(Asparouhov & Muthén, 2010; Holtmann et al., 2016).

베이저안 방법에서 사전분포를 설정할 때는 모수로 예상되는 값의 확률밀도를 크게 함으로써 연구자가 가진 사전 정보를 추정에 포함한다. 이러한 특징을 바탕으로 베이저안 추정에서는 분산, 공분산, 잔차분산 등의 분산 모수(variance parameter)에 음수가 아닌 범위에서만 정의되는 사전분포를 주로 지정해 빈도주의 추정에서 종종 나타나는 Heywood case 발생을 피할 수 있다. 분산 모수의 사전분포로는 감마(gamma)분포, 역감마(inverse gamma)분포, 위샛트(Wishart)분포 등이 사용되며, 이중 주로 쓰이는 분포는 역감마분포이다(Gelman, 2006). 역감마분포도 정규분포와 마찬가지로 공액의 특성을 갖고 있으며, $\Gamma^{-1}(\alpha, \beta)$ 로 표기하는데, α 는 형태 모수(shape parameter), β 는 척도 모수(scale parameter)이며, 두 하이퍼파라미터는 양수이다. 역감마분포는 정규분포처럼 직관적으로 하이퍼파라미터를 정하기 어렵지만, 모수에 대한 사전 지식이 있다면 식 6에 제공되는 최빈값 공식을 이용할 수 있다(McNeish, 2019).

$$\text{mode} = \frac{\beta}{\alpha + 1} \quad (6)$$

아닌 정확한 해를 구할 수 있다.

역감마분포는 정적 편포의 형태를 띠고 있으며, α 와 β 의 값이 증가할수록 분포의 값들이 0 근처로 모이고 분포의 꼬리는 얇아지는 경향을 보이므로 식 5를 통해 α 와 β 를 조절해 원하는 최빈값을 가지며 사전 정보에 대한 확신 정도를 적절히 반영하도록 분포의 범위 조절이 가능하다.

정보성

사전분포는 범위가 좁을수록, 즉 분산 하이퍼파라미터의 크기가 작을수록 추정에 미치는 영향이 커진다. 따라서 연구자는 사전 정보의 의도한 대로 추정에 영향을 미칠 수 있게 사전분포의 범위를 적절히 지정해야 한다(Depaoli, 2014; Muthén & Asparouhov, 2012). 예를 들어, 믿을만한 사전 정보를 가졌다면 범위가 좁은 사전분포를 지정해 사전분포가 추정에 미치는 영향을 높일 수 있고, 반대로 정보가 충분하지 않다면 넓은 범위의 사전분포를 지정해 사전분포가 추정에 미치는 영향을 줄일 수 있다. 이처럼 사전분포에 반영된 사전 정보의 정확도를 일컬어 사전분포의 정보성이라 한다. 분산 하이퍼파라미터의 크기를 통해 결정되는 사전분포의 정보성은 실제 연속적이지만, 대다수 연구자는 그림 1과 같이 무정보(noninformative), 약한 정보(weakly informative), 강한 정보(strong informative)로 나눈다(Gelman et al., 2013; Holtmann et al., 2016; McNeish, 2016). 많은 연구자가 이와 같은 용어를 사용하고 있기는 하지만, 사실 세 용어에 대한 해석은 그다지 명확하지 않다(van de Schoot et al., 2017). 예를 들어, Depaoli와 Clifton(2015) 및 Gelman(2006) 등은 무정보 사전분포를 ‘모수에 대해 연구자가 가진 정보가

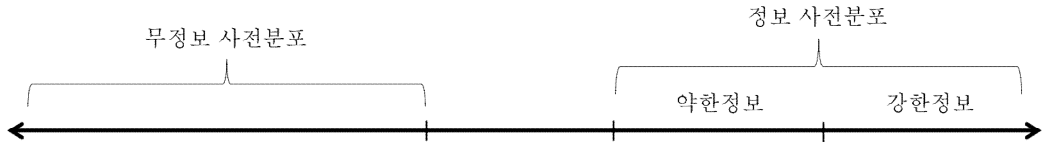


그림 1. 사전분포의 정보성

불확실할 때 사용하는 사전분포', 약한 정보 사전분포를 '무정보 사전분포보다는 정보가 많으나, 정확하지는 않은 정보를 가진 사전분포', 강한 정보 사전분포를 '믿을만한 정보를 포함한 사전분포'라고 설명하는데, 이와 같은 모호하고 부정확한 표현들이 연구자들에게 혼란을 야기하는 것이 현실이다. 사실 이런 이유로 그림 1에서도 어디부터 어디까지가 무정보 사전분포의 영역이고, 또 정보 사전분포의 영역인지 명확하게 표시되어 있지 않다. 더불어, 무정보 사전분포는 디폴트 사전분포와 혼용되는 경우가 잦는데, 이는 실제로 연구자들이 정보가 없는 모수에 무정보 사전분포보다 프로그램의 디폴트 사전분포를 지정하는 경우가 많기 때문이다(Fang et al., 2019; Hagger & Hamilton, 2018).

합리적인 값을 포함하도록 범위를 조정해 결정되는 사전분포의 정보성(Gelman et al., 2008)과 사전분포의 종류는 명확한 기준을 제시하기에 어려움이 있는 것이 어쩌면 당연하다. 하지만 베이지안 추정에서 올바르게 사전분포를 이해하고, 사전 정보를 잘 반영하기 위해서는 대략적인 가이드라인이 필요하다. 그림 2는 사전분포를 지정할 때 연구자의 의도와 연구자가 가진 사전 정보의 양에 따라 사전분포의 종류를 나누어 도식화한 것이다. 그림에서는 우선 연구자가 사전분포를 직접 지정하고자 하는지 아닌지를 첫 번째 분류 기준으로 두었다. 사전분포를 직접 지정하지 않는 경우 프로그램에서 제공하는 디폴트 사전분포를 선택할 수 있으며, 직접 지정하고자 한다면 가지고 있는 사전 정보의 유무와 그에

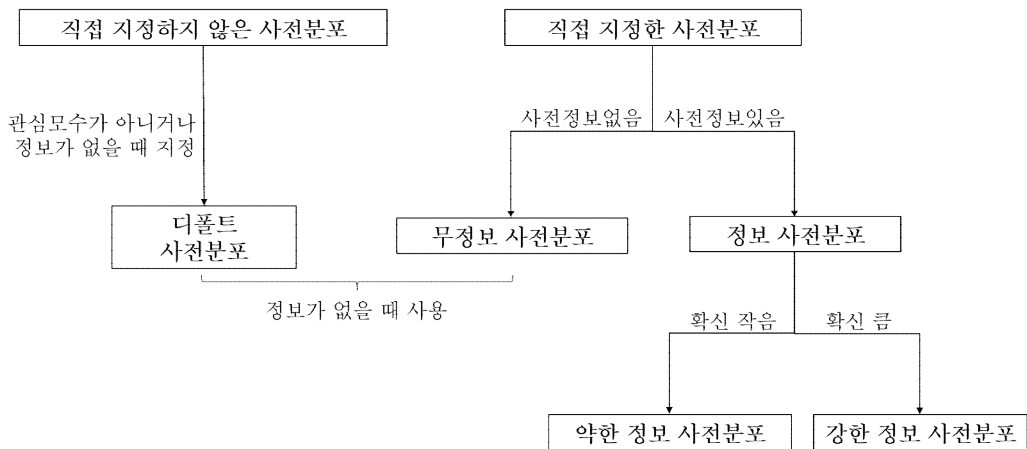


그림 2. 연구자의 의도와 사전 정보의 양에 따른 사전분포의 종류

대한 믿음의 정도에 따라 사전분포의 정보성이 정해진다.

사전분포를 직접 지정하지 않는 방법

디폴트 사전분포

현재 주요 프로그램에서 제공하는 디폴트 사전분포는 주로 관심이 적은 모수나 정보가 없는 모수의 사전분포로 사용되고 있다 (McNeish, 2016; van de Schoot et al., 2017). 디폴트 사전분포는 모든 모수에 사전분포를 지정해야 한다는 부담을 덜어줘 베이지안 분석에 대한 접근성을 높여주었으나(van Erp et al., 2018), 한편으로는 연구자들이 과도하게 디폴트 사전분포에 의존하고 있다는 점이 문제로 지적되기도 한다(Smid et al., 2020a). 이는 디폴트 사전분포가 특정한 상황에서 연구자의 의도와 다르게 작동하기도 하는 등(Gelman, 2006; McNeish, 2019) 추정에 오류를 일으킬 가능성이 있기 때문이다. 이에 연구자들이 많이 사용하는 베이지안 추정 소프트웨어에서

디폴트 사전분포로 어떤 분포가 제공되는지와 현재까지 알려진 디폴트 사전분포의 문제점에 대한 논의가 필요하다. 본 연구에서 디폴트 사전분포는 단변량 연속형 변수에 초점을 맞춰 정보를 제공한다.

디폴트 사전분포의 종류

현재 사회과학 분야에서 베이지안 분석을 위해 사용되는 대표적인 프로그램 중 디폴트 사전분포를 제공하는 프로그램은 표 1에 제공된 *Mplus*(Muthén & Muthén, 1998-2019), R(R Core Team, 2015)의 *blavaan* 패키지(Merkle & Rosseel, 2015), *WinBUGS*(Lunn et al., 2000)이다. *JAGS*(Plummer, 2003)와 *Stan*(Carpenter et al., 2017)도 많이 사용되는 베이지안 프로그램이지만, *JAGS*는 디폴트 사전분포를 제공하지 않고, *Stan*은 모든 모수의 디폴트 사전분포로써 균등분포 $U(-\infty, \infty)$ 를 제공하므로 논의에서는 제외한다. 또한 *WinBUGS* 메뉴얼에서 사용된 $\Gamma^{-1}(0.001, 0.001)$ 은 많은 연구자가 분산 모수의 사전분포로 사용하고 있기는 하나, 디폴트 사전분포가 아니므로 표에는 포함하지 않았다.

표 1. 베이지안 프로그램에서 제공하는 디폴트 사전분포의 종류

프로그램	사전분포(Priors)	
	위치 모수 (location parameter)	분산 모수 (variance parameter)
Mplus	$N(0, 10^{10})$	$\Gamma^{-1}(-1, 0)$
blavaan(R package)	$\nu \sim \text{dnorm}(0, 1e-3)$ $\alpha, \beta, \lambda \sim \text{dnorm}(0, 1e-2)$	$\theta, \psi \sim \Gamma^{-1}(1, 0.5)$
WinBUGS	$U[-\infty, \infty]$	

ν : Observed variable intercept, α : Latent variable intercept, λ : Loading, β : Regression intercept and slope, θ : Observed variable precision parameters, ψ : Latent variable precision parameters

*Mplus*는 위치 모수에 $N(0, 10^{10})$ 을 디폴트 사전분포로 제공한다. 이 분포는 매우 큰 분산으로 인해 어떤 부분을 확대해 보아도 평평한 모양을 띠는 특징을 갖는다. 이처럼 디폴트 사전분포는 사전분포가 가진 정보를 없애기 위해 매우 넓은 상수 범위에서 확률밀도를 갖는 분포로 지정되는 것이 대부분이다. 하지만 이로 인해 확률분포의 기본 성질 중 하나인 모든 사건이 일어날 확률을 더했을 때 1(또는 상수)이 되어야 하는 성질을 위배하는 경우가 종종 있다. 이런 분포를 부적절한 사전분포(improper priors)라 하는데, 부적절한 분포란 전 범위에 걸친 확률밀도함수를 적분했을 때 1이 되지 않는 분포를 뜻한다(Gelman et al., 2013; Yuan & Mackinnon, 2009). 부적절한 사전분포는 기술적인 문제로 인한 계산을 오류를 발생시킬 수도 있으므로 사용에 주의가 필요하다(Asparouhov & Muthén, 2010; van de Schoot et al., 2015; van Erp et al., 2018). *Mplus*의 디폴트 사전분포인 $N(0, 10^{10})$ 은 정규분포임에도 불구하고, 너무 큰 분산 하이퍼파라미터로 인해 성장혼합모형(growth mixture model) 추정에서 부적절한 사전분포로 기능해 큰 편향을 일으키는 문제가 보고되기도 한다(Depaoli, 2014). *blavaan*은 측정모형에서 절편 모수(ν)의 디폴트 사전분포로 $\text{dnorm}(0, 1e-3)$ 을 제공하고, 구조모형의 절편 모수, 요인부하, 잠재변수 간 직접 경로에는 $\text{dnorm}(0, 1e-2)$ 을 디폴트 사전분포로 제공한다. 이때 *blavaan* 디폴트 사전분포의 형태가 *Mplus*에서 제공되는 사전분포와 매우 다른 것을 볼 수 있는데, 이는 *blavaan*이 분산의 역수인 'precision'을 기본으로 사용하기 때문이다(Plummer, 2003; Spiegelhalter et al., 2003). 'precision'은 그 값이 커질수록 사전분포가 추

정에 미치는 영향도 커지므로 직관적으로 이해하기 쉽다는 장점이 있어 베이지안 방법에서 자주 사용된다(van de Schoot et al., 2014). WinBUGS는 $[-\infty, \infty]$ 범위에 균등분포를 지정할 수 있는 'dflat()' 코드를 제공한다(Lunn et al., 2012). $U[-\infty, \infty]$ 는 모수가 될 수 있는 모든 범위의 값이 갖는 확률밀도가 같으므로 추정에 미치는 영향이 없는 것으로 간주하지만, 이 역시도 부적절한 사전분포로 기능할 수 있다.

분산 모수의 디폴트 사전분포로는 주로 역감마분포가 이용된다. *Mplus*는 $\Gamma^{-1}(-1, 0)$ 을 디폴트 사전분포로 제공한다(Asparouhov & Muthén, 2010). 역감마분포의 하이퍼파라미터는 모두 양수로 지정되어야 하지만, *Mplus*는 새롭게 정의한 $\Gamma^{-1}(-1, 0)$ 을 디폴트 사전분포로 제공한다. $\Gamma^{-1}(-1, 0)$ 의 실제 분포 형태는 $U[0, \infty)$ 이고 그림 3과 같다. $\Gamma^{-1}(-1, 0)$ 은 음수 범위에서 확률밀도를 갖지 않아 추정치로 음수가 나올 가능성이 없고, 0 이상의 모든 값은 균등한 확률밀도를 갖는다. *blavaan*은 분산 및 공분산에 디폴트 사전분포로 $\Gamma^{-1}(1, 0.5)$ 를 제공하며 형태는 그림 4와 같고, 이는 대부분의 양수 범위에서 거의 정보를 가지고 있지 않다. Merkle과 Rosseel(2015)은 분산 모수의 디폴트 사전분포를 설정할 때 Barnard et al. (2000) 및 Muthén과 Asparouhov(2012)를 참고하여 설정했으나, *Mplus*와 달리 추정에 문제가 없는 적절한 사전분포만을 제공한다고 밝혔다. 공식적으로 WinBUGS는 분산 모수에 제공하는 디폴트 사전분포가 없으나, 프로그램 메뉴얼에서 사용된 $\Gamma^{-1}(0.001, 0.001)$ 가 디폴트 사전분포처럼 사용되며, 이는 그림 5와 같다. $\Gamma^{-1}(0.001, 0.001)$ 은 Spiegelhalter et al. (2003)

이 Jeffreys(1946) 방법에 따라 계산한 무정보 사전분포이다. Jeffreys(1946)는 무정보 사전분포라면 분포의 식을 다른 매개변수로 치환하여도 여전히 정보가 없는 형태를 가져야 한다는 개념을 바탕으로 사전분포를 고안하였다.

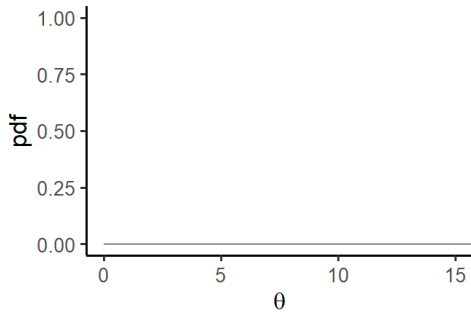


그림 3. Mplus의 $\Gamma^{-1}(-1, 0)$

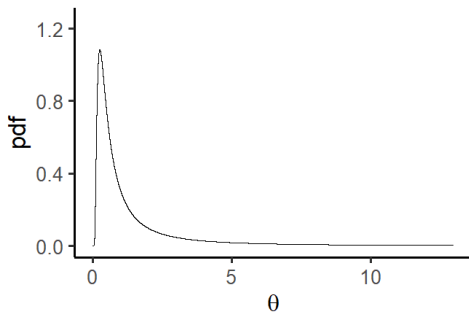


그림 4. blavaan의 $\Gamma^{-1}(1, 0.5)$

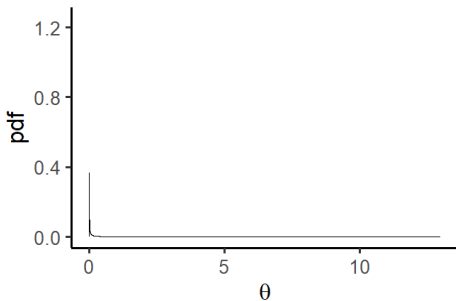


그림 5. WinBUGS의 $\Gamma^{-1}(0.001, 0.001)$

디폴트 사전분포의 문제점

디폴트 사전분포는 편리한 도구지만 예상치 못한 문제들을 유발할 수 있다. 예를 들어, 무정보 사전분포를 의도하였으나 정보 사전분포로 기능하기도 하고, 부정확한 추정치를 제공하기도 한다. 이러한 문제로 인해 몇몇 연구는 디폴트 사전분포의 무분별한 사용에 대해 우려를 표하며 디폴트 사전분포가 연구자의 의도대로 작동하지 않을 수 있음을 지적한다 (Gelman et al., 2013; McNeish, 2019; Smid & Winter, 2020). 디폴트 사전분포가 실제 연구에서 발생시키는 구체적인 문제들은 다음과 같다.

첫째, 작은 표본크기 연구에서 모수의 종류와 상관없이 추정 결과에 문제가 발생할 확률이 높다. 일반적으로 연구자들은 작은 표본크기를 가진 연구에서 베이지안 방법을 빈도주의 방법의 대안으로 기대하지만(McNeish, 2016), 베이지안 방법을 이용한다는 사실만으로 작은 표본크기 문제가 해결되지는 않는다. 작은 표본크기로도 타당한 결과를 얻는 것은 적절한 사전분포를 지정했을 때 가능하다(Smid et al., 2020a). 실제로 디폴트 사전분포는 매우 큰 분산으로 인해 작은 표본크기를 가진 분석에서 정상적으로 작동하지 않는 문제(예, 낮은 검정력, 크게 편향된 추정치, 다른 방법에 비해 큰 표준 오차 등)가 보고된다(Hox, 2020; McNeish, 2016; van Erp et al., 2018). 즉, 표본크기가 작은 경우 넓은 범위에 걸친 디폴트 사전분포를 이용하면 적절한 사후분포를 얻는 데 어려움을 겪을 수 있으며(Muthén & Asparouhov, 2012), 이러한 문제는 분산 모수에서 더욱 두드러진다(Smid et al., 2020a).

둘째, 부적절한 사전분포를 사용하면 부적

절한 사후분포가 발생할 수 있다(Asparouhov & Muthén, 2010; Browne & Draper, 2006). 앞서 소개한 사전분포 중에서는 WinBUGS의 $U[-\infty, \infty]$ 와 *Mplus*의 $\Gamma^{-1}(-1, 0)$ 이 부적절한 사전분포이다. 또한 $\Gamma^{-1}(0.001, 0.001)$ 도 부적절한 사전분포에 속한다. 이 밖에도 분포의 범위가 매우 넓게 지정된 분포는 상황에 따라 부적절한 사전분포로 기능할 수 있다. 추정에서 부적절한 사후분포의 발생이란 곧 추정의 실패를 의미한다. 따라서 적절한 사전분포의 사용은 다수의 연구자가 권고하는 사항이다(Hobert & Casella, 1996; van de Schoot et al., 2015). 특히 Gelman(2006)은 분산 모수의 사전분포로 역감마분포를 사용할 때, 역감마분포의 하이퍼파라미터를 무조건 작은 값으로 설정하면 부적절한 분포가 될 확률이 높으므로 합리적인 값을 설정해야 한다고 강조하였다.

셋째, 무정보 사전분포로 기능할 것이라 예상하고 사용한 디폴트 사전분포가 상황에 따라 정보 사전분포로 기능하는 것이다(McNeish, 2016; Smid & Winter, 2020). Gelman(2006)은 다층모형 연구에서 실제로 거의 0에 가까운 학교 간 분산의 추정을 위한 무정보 사전분포로 $U(0, \infty)$, $\Gamma^{-1}(1, 1)$, $\Gamma^{-1}(0.001, 0.001)$ 을 설정하고 결과를 비교하였다. 균등분포를 설정한 추정 결과에는 우도함수가 많이 반영되어 0 근처의 값이 큰 확률밀도를 갖는 사후분포가 적절하게 도출되었다. 반면에 역감마분포를 사전분포로 지정한 추정에서는 자료의 정보가 추정치에 제대로 반영되지 않고, 모두 0에서 벗어난 값이 큰 확률밀도를 갖는 사후분포가 추정되었다. 균등분포의 결과와 달리 역감마분포를 사용한 추정에서는 설정된 사전분포와 비슷한 형태의 사후분포가 나타났던

것이다. 이는 역감마분포가 0 근처에서 급격한 값의 변화를 보이고, 0에 아주 가까운 값들은 매우 낮은 확률밀도를 가진 형태를 띠기 때문이다. 또한 Depaoli(2012)는 blavaan에서 구조모형의 위치 모수에 제공되는 디폴트 사전분포인 $N(0, 100)$ 을 약한 정보 사전분포로 규정하고 연구를 진행하였다. 이는 디폴트 사전분포로 제공되는 분포일지라도 항상 무정보 사전분포가 되는 것은 아님을 보여주는 단적인 예이다.

상기한 문제들은 모두 디폴트 사전분포 사용으로 인해 발생할 수 있는 추정의 실패나 잘못된 추정에 대한 것이다. Depaoli와 Clifton(2015)은 정보가 없는 모수일지라도 더 나은 추정을 위해서는 모든 사전분포의 지정에 신중한 고민이 필요하다고 강조한다. 특히 작은 표본크기의 연구는 사전분포에 무척 민감하기 때문에, 확산분포(diffuse priors)나 디폴트 사전분포는 빈도주의 방법보다 더 편향된 결과를 발생시킬 수 있다(McNeish, 2019). 따라서 연구자는 사전 정보가 없을 때 디폴트 사전분포 대신 무정보 사전분포를 지정하는 것을 고려해야 한다.

사전분포를 직접 지정하는 방법

직접 사전분포를 지정하는 것은 약간의 어려움이 따르지만, 디폴트 사전분포로 발생하는 문제점들을 해결하고, 더 나은 추정 결과를 얻을 수 있다(Smid et al., 2020a). 지금부터 연구자에 의해 직접 지정되는 무정보 사전분포와 정보 사전분포를 다룬다. 먼저 무정보 사전분포의 필요성과 선행연구들에서 제안된 무정보 사전분포를 지정하는 방법에 대해 알

아보고, 정보 사전분포를 지정할 때 참고할 수 있는 연구들을 요약해 제공한 뒤, 많은 연구에서 지정된 사전분포를 취합해 정보성 선택에 참고할 수 있도록 사전분포 정보성 지정의 경향성을 탐색한다.

무정보 사전분포

베이저안 분석을 이용하려면 모든 모수에 사전분포를 할당해야 하지만 연구자가 모든 모수에 대한 정보나 사전 지식을 찾기란 매우 어렵다(Smid & Winter, 2020). 따라서 모수에 대한 정보가 없을 때, 추론에 영향을 거의 미치지 않으면서도 추정에서 최소한의 역할을 할 수 있는 분포가 필요하다(Gelman et al., 2013). 이러한 분포를 무정보 사전분포라고 하며, 가장 간단한 예는 균등분포나 분산이 큰 정규분포이다(Muthén & Asparouhov, 2012). 이외에도 무정보 사전분포를 설정하기 위해 제안된 방법들이 있으며, 본 연구에서는 이에 대해 자세히 논의한다.

무정보 사전분포를 지정하는 방법

첫 번째 방법은 경험적 베이지 사전분포(empirical Bayes priors, EB priors)를 사용하는 것이다. EB 사전분포는 수집한 자료의 정보를 바탕으로 설정된 사전분포이다. 이론적으로 수집된 표본의 정보는 사전분포 설정에 사용될 수 없는데, 이는 더블디핑으로 인하여 사후분포의 분산이 과소추정되는 문제를 발생시키기 때문이다. 그러나 모수에 대한 정보를 얻기 어려운 경우, 더블디핑으로 인한 문제는 피하면서 수집된 자료를 사전분포의 지정에 이용할 수 있는 몇 가지 방법이 있다. 본 연구에서는 van Erp et al. (2018)이 제안한 방법

을 소개한다. 먼저 위치 모수에 대해 단순화된 모형 식 7을 설정한다.

$$y_i = \alpha + \epsilon_i, \quad (i = 1, 2, \dots, n),$$

$$\epsilon_i \sim N(0, \sigma^2) \quad (7)$$

위에서 y_i 는 자료, α 는 위치 모수, ϵ_i 는 오차, σ^2 은 알려지지 않은 분산이다. 다음으로 α 에 대한 사전분포를 설정한다($\alpha \sim N(\mu_\alpha, \tau_\alpha^2)$). 분포의 평균은 사전 정보를 기반으로 정해야 하지만, 무정보 사전분포를 설정할 때는 참고할 정보가 없으므로 임의로 0을 할당한다. 분산 하이퍼파라미터인 τ_α^2 은 $\frac{1}{n} \sum_{i=1}^n y_i^2$ 으로 구할 수 있고, 이는 곧 $E[Y^2]$ 과 같다. $E[Y^2] = Var(Y) + (E[Y])^2 = \sigma^2 + \mu_\alpha^2$ 이므로, $\hat{\tau}_\alpha^2 = \hat{\sigma}^2 + \hat{\mu}_\alpha^2$ 이 된다. $\hat{\sigma}^2$, $\hat{\mu}_\alpha^2$ 에는 수집한 자료로부터 ML을 통해 얻은 추정치를 대입한다. 이와 같은 방법은 하나의 자료를 동일한 모형에 두 번 사용한 것은 아니므로 데이터의 이중사용 문제는 피하면서도 타당한 사전분포 설정을 돕는 방법으로 알려져 있다.

두 번째 방법은 표본크기가 어느 정도 확보된 경우에 데이터를 무작위로 두 세트로 나누어 한 세트를 분석해 나온 추정치를 다른 세트 분석의 사전분포 설정에 이용하는 방법이다. 이 방법을 적용한 예로는 교통안전 성능기능(traffic safety performance functions, SPF)을 개발하기 위한 연구가 있다. Yu와 Abdel-Aty (2013)는 251개의 자동차 충돌 자료를 무작위로 나누어 하나의 세트는 무정보 사전분포를 이용하여 모형을 추정하였다. 남은 데이터 세트는 이 결과를 바탕으로 지정한 정보 사전분포를 통해 실제 SPF를 개발하기 위한 추정에

사용되었다. Yu와 Abdel-Aty(2013)는 정보 사전 분포를 지정하는 것이 데이터를 나누지 않고 디폴트 사전분포로 추정하는 것보다 더 정확한 결과를 산출할 수 있었기 때문에 이 방법을 채택했다고 밝혔다.

이밖에 제안된 방법으로는 추정치가 될 만한 값의 범위에서만 정의되도록 사전분포를 한정하는 방법이 있다(Lunn et al., 2012; Miočević et al., 2017). 예를 들어, 5점 Likert 척도를 사용한 모수의 평균을 구한다면 그림 6과 같이 사전분포를 1부터 5의 범위에서만 정의된 균등분포로 지정한다(van de Schoot et al., 2014). 또는 기본적인 배경지식을 바탕으로 값이 분포되어 있을 범위를 고려하는 정규분포를 설정한다(Zondervan-Zwijnenberg et al., 2017). 또한 참고할 선행연구가 있더라도 정보에 대한 확신이 부족하다면(예, 다른 척도 사용, 연구 대상 집단의 불일치 등), 이 정보가 분석에 미치는 영향을 줄이고 싶을 수도 있다. 이때는 선행연구의 추정치를 바탕으로 사전분포의 평균 하이퍼파라미터를 지정한 뒤, 분산 하이퍼파라미터는 크게 설정하는 것을 고려해 볼 수 있다(Cortopassi et al., 2017).

앞서 언급한 방법들은 모든 모수에 적용 가능한 방법이었다. 다음은 분산 모수의 무정보

사전분포 설정을 위해 제안된 방법들을 알아 본다. 첫째, Gelman(2006)은 다층모형을 이용한 연구에서 집단 간 분산 모수의 사전분포로 Half-Cauchy 분포를 사용할 것을 추천하였다. Half-Cauchy 분포(그림 7)는 자유도가 1인 t 분포를 양수 범위에서만 정의한 분포이다. 분산 모수의 사전분포로 주로 사용되는 역감마분포는 원점을 지나야 한다는 특징으로 인해 0 근처에서 급격한 확률밀도의 변화를 보이며 (Polson & Scott, 2012), 0에 가까운 작은 값을 갖는 모수에 사용했을 때는 분석 결과에 문제가 일어날 확률이 높아진다(Gelman et al., 2013). 이에 분산 모수의 값이 0에 가까울 것으로 예상된다면, 0 근처에서 큰 확률밀도를 가지면서 완만한 형태를 띠고 있는 Half-Cauchy 분포가 추천된다. Gelman(2006)은 Half-Cauchy 분포를 사용할 때, 분포가 갖는 정보를 줄이기 위해 분산으로 예상되는 값을 포함하면서 분포가 넓은 범위에 걸쳐 척도(scale)를 설정하였다. 척도는 Half-Cauchy 분포의 분산을 의미하며, 척도가 커질수록 분포의 폭이 넓어진다. McNeish와 Stapleton(2016)은 시뮬레이션 연구를 통해 작은 표본크기를 가진 다층모형 연구에서도 Half-Cauchy 분포가 잘 기능함을 밝혔으며, Polson과 Scott(2012)은

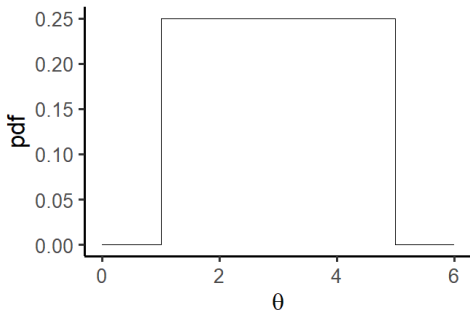


그림 6. 균등분포 $U[1, 5]$

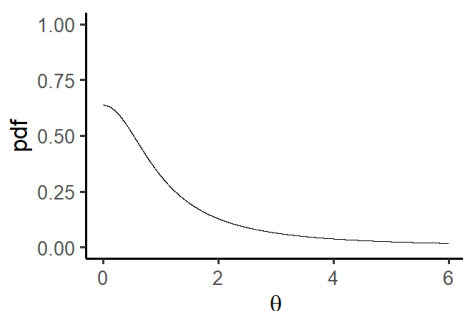


그림 7. Half-Cauchy 분포

Half-Cauchy 분포가 다층모형의 분산 모수(scale parameter)에 사용될 수 있는 사전분포의 종류 중 가장 합리적인 분포라고 하였다.

둘째는 선행연구를 바탕으로 역감마분포의 하이퍼파라미터를 계산할 수 있는 방법이다. 분산 모수는 추정과 해석이 까다롭기 때문에 (van Erp et al., 2018) 참고할 선행연구가 있다면 하이퍼파라미터를 계산할 수 있는 식을 이용해 적당한 분포를 탐색하는 것이 제안된다. Gelman et al. (2013)은 선행연구의 자유도와 분산 모수의 추정치를 통해 역감마분포의 하이퍼파라미터를 계산할 수 있는 식을 제안했다. 역카이제곱분포는 역감마분포의 특수한 형태이므로 식 8과 같은 유도가 가능하다.

$$\alpha = \frac{\nu}{2}, \beta = \alpha \times s^2 \quad (8)$$

위에서 ν 는 역카이제곱분포의 자유도를 의미하며, s^2 은 선행연구의 분산 모수 추정치이다. 만약 선행연구의 표본크기가 30이고, 분산 추정치가 0.3이라면 $\alpha = 30/2 = 15$, $\beta = 15 \times 0.3 = 0.45$ 를 통해 모수의 사전분포가 $\Gamma^{-1}(15, 0.45)$ 로 지정된다. 이때 만약 선행연구의 표본크기가 크다면, 하이퍼파라미터 값이 커지므로 분포의 폭이 좁아지게 된다. 좁은 범위의 분포는 무정보 사전분포를 지정하고자 하는 의도와 맞지 않으므로, 표본크기(ν)를 더 작은 값으로 대체하여 사전분포의 불확실성을 높일 수 있다(Smid et al., 2020a).

정보 사전분포

정보 사전분포란 모수에 대한 사전 지식과 이에 대한 확신의 정도를 반영해 지정된 분포

이다. 특히 모수에 대한 옳은 정보를 바탕으로 적절한 정보성이 할당된 사전분포를 정밀한 정보 사전분포(accurate informative prior)라 하며 이는 정확한 추정을 돕는다(McNeish, 2016; Smid et al., 2020a). 정밀한 정보 사전분포를 지정하기 위해서는 사전 정보를 탐색하고, 다른 연구들이 사전분포를 어떻게 지정하는지 알아볼 필요가 있다. 이를 위해 본 연구에서는 정보 사전분포를 지정한 경험적 선행연구들을 제공하고, 메타 분석을 활용한 연구를 확인한다. 다음으로는 사전분포 지정과 관련하여 방법론 연구에서 제공된 예시를 다룬다. 이후 현재까지 출판된 다수의 베이زي안 경험적 연구 및 시뮬레이션 연구에서 직접 지정된 사전분포의 정보성 기준을 탐색하고 경향에 대해 논의한다.

정보 사전분포를 지정한 연구

Cortopassi et al. (2017)은 남성 성 소수자를 대상으로 나이, 인종, 수입, 자기 은폐(self-concealment)가 인지적 재해석(cognitive reappraisal)을 통하여 약물 의존에 미치는 영향을 밝히기 위해 그림 8과 같은 경로모형을 설정하였다. Cortopassi et al. (2017)은 사전분포에 대한 정보를 얻기 위해 동일한 척도를 사용한 선행연구를 참고하였다. 먼저, 처방약을 이용하는 개인의 약물 의존도를 연구한 Kelly et al. (2015a)의 약물 의존도 결과($M = 2.25$, $SD = 2.42$)를 바탕으로 약물 의존 평균에 대한 사전분포는 $N(2.25, 100)$ 으로 설정하였다. 다음으로 인지적 재해석 평균의 사전분포 지정을 위해 Mohajerin et al. (2013)을 참고하였다. Mohajerin et al. (2013)은 인지적 재해석 하위 척도를 사용하여 측정한 단일 물질에 의존하고 있는 개인의 인지적 재해석 점수

($M = 21.93$, $SD = 5.75$)와 여러 물질에 의존하고 있는 개인의 인지적 재해석 점수 ($M = 23.57$, $SD = 7.71$)에 대해 연구하였고, 이를 바탕으로 Cortopassi et al. (2017)은 인지적 재해석 평균의 사전분포를 $N(22.0, 100)$ 으로 설정하였다. 또한 Kelly et al. (2015b)에서 밝힌 나이가 약물 의존에 미치는 영향 ($B = -0.02$)과 인종이 약물 의존에 미치는 영향 ($B = -0.04$)을 통해 회귀계수의 사전분포는 각각 $N(-0.02, 100)$, $N(-0.04, 100)$ 으로 지정하였다. Cortopassi et al. (2017)은 다른 연구의 추정치가 갖는 불확실성을 반영하기 위해 모든 분포의 분산을 100으로 설정하였고, 정보를 찾지 못한 나머지 모수에는 *Mplus* 디폴트 사전분포를 할당하였다.

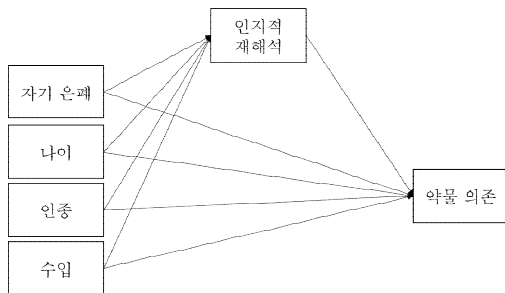


그림 8. 정보 사전분포를 설정한 경로모형

선행연구를 참고한 또 다른 예는 Falkenström et al. (2015b)의 작업 동맹(working alliance)에 대한 확인적 요인분석 연구이다. Falkenström et al. (2015b)은 총 4개의 모델을 상정하였는데, 각각은 단일 요인 모형, 2요인 모형, 요인 간 상관이 없는 쌍요인(bifactor) 모형, 요인 간 상관이 있는 쌍요인 모형이다. 먼저 선행연구의 2요인 모형 공유분산이 50%임을 바탕으로 단일 요인 모형의 요인부하 및 쌍요인 모형의 일반 요인의 요인부하에 $N(0.7, 0.02)$ 를 사

전분포로 할당하였다. 평균 하이퍼파라미터는 제공하여 0.5에 가까운 값($0.7^2 = 0.49 \approx 0.50$)을 선택한 결과이다. 다음으로 Falkenström et al. (2015b)은 2요인 모형과 쌍요인 모형의 Bond 및 GoalTask 요인의 부하가 일반 요인보다 작은 부하를 가질 것이라는 믿음을 바탕으로 요인부하 사전분포를 $N(0.5, 0.02)$ 로 설정하였다. 또한 모든 사전분포의 분산 하이퍼파라미터를 설정할 때는 효율적인 추정이 가능한 몇 가지 분산을 추린 뒤, 실험을 통해 그중 가장 큰 값을 선택함으로써 사전분포의 정보성을 낮추고자 하였다. 그 결과 상기한 요인부하 사전분포의 분산 하이퍼파라미터는 0.02로 지정하였다. 다음으로 교차 요인부하의 사전분포는 Muthén과 Asparouhov(2012)의 연구를 참고해 모두 $N(0, 0.01)$ 로 지정하였다. 마지막으로, 잠재평균과 절편의 평균 하이퍼파라미터는 저자들의 선행연구인 Falkenström et al. (2015a)의 결과와 동일하게 지정하였으며, 분산 하이퍼파라미터에는 요인부하 사전분포와 마찬가지로 실험 후 0.5를 할당하였다. 따라서 잠재평균에는 $N(0, 0.5)$, 절편에는 $N(5.5, 0.5)$ 가 지정되었다. 저자들은 분산 하이퍼파라미터를 설정할 때, 모수가 될 가능성이 있는 값이 사전분포에 모두 포함되도록 지정하였으며 디폴트 사전분포는 사용하지 않았다.

Xu et al. (2015)은 베이지안 분석을 위한 사전 정보 탐색에 메타 분석을 활용하였다. 연구의 목적은 무정보 사전분포와 세 종류의 메타 분석 결과를 바탕으로 지정한 사전분포로부터 도출된 결과를 비교함으로써 제한된 자료로부터 가장 나은 고속도로 사고 위험 예측 모형을 개발하는 것이었다. Xu et al. (2015)은 총 36개 논문에 실린 실시간 교통 관련 연구

의 결과를 수집해 세 종류의 메타분석(고정효과, 무선효과, 메타 회귀분석)을 설계하였다. 모형 1-3은 순서대로 통합 효과(overall pooled effect)만 측정하는 고정효과 메타 분석 결과, 연구 간의 이질성과 변동성도 함께 추정하는 무선효과 메타 분석 결과, 연구 결과 간 체계적인 차이를 추가로 분석하는 메타 회귀분석 결과를 이용해 사전분포를 설정한 모형이고, 모형 0은 위치 모수에 $N(0, 10^6)$, 분산 모수에 $\Gamma^{-1}(0.001, 0.001)$ 을 무정보 사전분포로 이용한 모형이다. 이들은 분석 후 모형 0의 결과와 모형 1-3의 결과를 비교함으로써 가장 나은 예측 결과를 보이는 모형을 찾고자 하였다. 최종적으로 가장 좋은 결과는 메타 회귀분석을 이용하였을 때였으며, 이는 무정보 사전분포를 이용한 결과보다 충돌 예측 정확도를 평균적으로 15.1%까지 증가시키는 것으로 나타났다.

마지막으로 Zondervan-Zwijenburg et al. (2017)은 전문가의 의견을 사전분포 지정에 반영하였다. 예시는 대마초를 많이 피우는 학생과 피우지 않는 학생 간의 인지능력을 비교하는 실험연구이다. 6개월 간격으로 2년간 수집한 자료를 통해 잠재성장모형을 추정하였으며, 전문가에게 의견을 구하기 전에 선행연구와 문헌 연구 등을 통해 사전 정보를 찾았다. 사전분포 설정을 위해 6개의 메타 분석 연구를 검색하여 찾았으나, 연구 대상과 방법이 맞지 않아 참고하지 않았다. 다음으로 리뷰 논문을 검색해 9개의 연구를 찾았고, 이로부터 관심 모수인 성장률 차이(growth rate difference)에 대한 정보를 파악하였다. 다음으로 절편(intercept), 1차 기울기(linear slope), 2차 기울기(quadratic slope)의 값에 대한 정보를 얻기 위해 검색한 693개의 연구 중 다섯 가지의

기준을 통해 참고할 수 있는 13개의 연구를 추려내었고, 이 선행연구들을 바탕으로 사전분포를 설정하였다. 이후 설정한 사전분포가 대마초를 많이 피우는 집단에 편향적으로 설정되지 않았는지 확인하기 위해 전문가의 도움을 받았다. 자문을 구한 전문가는 발달 정신병리학 교수 한 명과 임상 심리학 한 명이었으며, 실험에 대한 정보는 알리지 않은 채로 설문을 진행하고 참고하였음을 밝혔다.

정보 사전분포의 정보성 기준

정밀한 사전분포를 설정하기 위해서는 정확한 평균 하이퍼파라미터뿐만 아니라, 분산 하이퍼파라미터를 적절한 값으로 할당하는 것이 중요하다. 분산 하이퍼파라미터는 사전분포의 정보성에 영향을 미치는데, 연구마다 임의의 기준으로 사전분포의 정보성을 구분한 채 연구를 수행하기 때문에 정보성을 나누는 기준이 모호한 것이 현실이다(van de Schoot et al., 2017). 이렇게 혼재된 기준은 정밀한 사전분포의 설정을 방해하며 잘못된 추정치를 산출하는 등 결과에 부정적 영향을 끼칠 수 있다(Depaoli, 2014). 예를 들어, 정확하지 않은 사전 정보를 바탕으로 강한 정보 사전분포를 설정한다면, 사전분포와 실제 모수의 분포 간 발생하는 차이로 인해 추정에 실패할 확률이 높아진다. 따라서 명확하지 않은 정보를 가졌을 때는 강한 정보 사전분포를 선택하기보다는 약한 정보 사전분포를 설정하는 것이 낫다(Depaoli, 2014; Miočević et al., 2020). 이처럼 가지고 있는 사전분포에 대한 정보를 분석에 적절히 반영하기 위해서는 정보성에 대한 판단과 연구자가 참고할 수 있는 지침이 필요하다. 이에 본 연구에서는 사전분포를 직접 설정한 이전 연구들을 취합해 통합 정리함으로써

써 정보성 기준의 경향을 파악하였다. 이를 위해 먼저 다수의 직접 지정된 사전분포를 수집했다. 정보 사전분포는 대부분 위치 모수에 할당되어있기 때문에, 본 연구는 연속형 변수이며 정규분포를 사전분포로 사용한 위치 모수에 초점을 맞추었다. 즉, 이전 연구 중에서 정보 사전분포로 범위를 한정하여 정보성을 구분한 경험적 연구 및 시뮬레이션 연구를 탐색하였다. 조건에 부합하는 15개의 연구에서 취합한 사전분포는 약 80개였고, 모형이나 모수의 종류는 구분하지 않았다. 그리고 본 연구에서는 사전분포의 평균 하이퍼파라미터에 대한 분산 하이퍼파라미터의 비(ratio)를 바탕으로 경향성을 찾았다. 비를 사용한 이유는 변수마다 단위가 달라 발생할 수 있는 혼란을 피하기 위함이며, 정리된 내용은 표 2에 제시된다.

수집한 각 연구에서 무정보 사전분포와 정보 사전분포는 모두 구분하고 있었다. 정보 사전분포는 세분하지 않는 경우도 있었으나,

대다수는 약한 정보 사전분포와 정보 사전분포로 구분하거나(Depaoli & Clifton, 2015; Muthen & Asparauhov, 2012), 약한 정보 사전분포와 강한 정보 사전분포로 나누었다(Holtmann et al. 2016; Smid et al., 2020b; van de Schoot, et al. 2015). 본 연구에서는 이를 모두 반영하여 무정보 사전분포와 정보 사전분포를 먼저 구분하고, 정보 사전분포를 다시 약한 정보 사전분포와 강한 정보 사전분포로 나누었다.

구분을 위해 수집한 사전분포의 평균 하이퍼파라미터에 대한 분산 하이퍼파라미터의 비를 모두 계산해 각 기준에 속하는 사전분포의 비율(proportion)을 확인하였다. 그 결과, 무정보 사전분포는 평균을 0으로 설정한 경우가 많았으나, 비를 계산할 수 있는 모든 무정보 사전분포는 평균 하이퍼파라미터에 대한 분산 하이퍼파라미터의 비가 100% 이상이였다. 정보 사전분포로 명명된 사전분포 중 96%는 평균 하이퍼파라미터에 대한 분산 하이퍼파라미

표 2. 정보 사전분포의 정보성 경향

사전분포의 종류	분산/평균 비(ratio) $\left(\left \frac{\sigma^2}{\mu} \right \times 100(\%) \right)$	비율(proportion)
무정보 사전분포 (noninformative prior)	1000% 초과	100%
	1000% 이하	0%
정보 사전분포 (informative prior)	100% 이상	4%
	100% 이하	96%
- 약한 정보 사전분포 (weakly informative prior)	100% 초과	11%
	10% 초과 100% 이하	67%
	10% 이하	22%
- 강한 정보 사전분포 (strong informative prior)	10% 초과 100% 이하	15%
	10% 이하	85%

터의 비가 100% 이하로 나타났다. 이중 약한 정보 사전분포로 명명된 사전분포들을 먼저 보면, 10% 초과 100% 이하의 범위에 67%가 포함되어 있었고, 강한 정보 사전분포로 명명된 사전분포들 중 85%는 분산 하이퍼파라미터의 크기가 평균 하이퍼파라미터 값의 10% 이하인 것으로 확인되었다. 종합해보면, 현재까지 사전분포를 직접 지정한 연구에서는 무정보 사전분포를 설정할 때 분산 하이퍼파라미터 값을 평균 하이퍼파라미터 값의 10배 이상으로 지정하였으며, 정보 사전분포는 분포의 분산 하이퍼파라미터가 평균 하이퍼파라미터의 100% 이내에서 설정된 경우가 대부분이었다. 또한, 정보 사전분포를 세분하는 연구자들은 분산 하이퍼파라미터가 평균 하이퍼파라미터 값의 10% 이하일 때 이를 강한 정보 사전분포로 판단하는 것으로 확인되었다.

경험적 적용 예시

앞서 다룬 사전분포를 어떻게 직접 지정할 수 있는지 실제 매개효과 예시에 적용하여 제공한다. 매개효과는 1900년대 초부터 연구되었으나 특히 Baron과 Kenny(1986) 이래로 많은 관심을 받아왔으며, 현재는 사회과학 분야 연구에서 가장 많이 사용되는 연구모형 중 하나이다(Preacher et al., 2007). 본 연구에서는 간단한 예시를 위해 그림 9와 같은 경로 매개모형(path mediation model)을 이용해 모형의 모든 모수에 사전분포를 직접 설정하는 방법을 보인다. 무정보 사전분포는 본문에 제안된 방법들을 바탕으로 지정하고, 정보 사전분포는 표 2에서 확인한 경향성 기준을 참고해 설정한다. 먼저 매개모형의 모수들에 대해 알아보고,

분석에 사용된 자료에 대한 간략한 정보를 제공한 뒤 사전분포 지정의 예시를 제공한다.

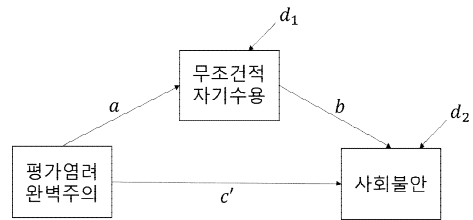


그림 9. 예시를 위한 연구모형

자료 및 연구 모형

분석에 사용된 자료는 378명의 대학생을 대상으로 수집되었다(Lee & Kim, 2018). 연구에 사용된 경로 매개모형은 세 개의 변수(독립변수, 매개변수, 종속변수)로 구성되어있고, 식 9~10과 같은 두 개의 회귀식으로 표현할 수 있다.

$$M = i_1 + aX + e_1 \quad (9)$$

$$Y = i_2 + c'X + bM + e_2 \quad (10)$$

위에서 i 는 회귀식의 절편을 의미하고, e 는 오차이다. 회귀계수 a 는 독립변수(X)가 매개변수(M)에 미치는 영향을 의미한다. c' 은 매개변수의 영향이 통제되었을 때 독립변수가 종속변수(Y)에 미치는 영향이고, b 는 독립변수를 통제하였을 때 매개변수가 종속변수에 미치는 영향이다. 설정된 가설은 평가염려 완벽주의(X)가 사회불안(Y)에 미치는 영향을 무조건적 자기수용(M)이 매개한다는 것이다. 본 예제의 모든 분석에는 *Mplus* 8.4를 이용하였다.

사전분포 설정

그림 9의 매개모형에서 사전분포를 지정할 모수는 경로계수(a, b, c')와 절편(i_1, i_2) 및 설명오차 (d_1, d_2)의 분산($\sigma_{d_1}^2, \sigma_{d_2}^2$)이다. 위치 모수인 경로계수에는 정규분포, 절편에는 균등분포, 분산 모수에는 역감마분포를 사용하였다. 사전분포 지정을 위해 Google Scholar에서 perfectionism, unconditional self-acceptance, social anxiety를 함께 제시어로 검색하였다. 검색 결과, 리뷰 논문이나 메타 분석을 진행한 연구는 없었고, 대학생을 대상으로 같은 변수나 비슷한 변수를 이용해 매개모형을 분석한 네 개의 경험적 연구를 찾을 수 있었으며, 이를 참고하여 사전분포를 지정하였다.

무정보 사전분포 설정

본 예시에서 무정보 사전분포를 설정할 모수는 절편(i_1, i_2)과 분산($\sigma_{d_1}^2, \sigma_{d_2}^2$)이다. 참고한 네 개의 경험적 연구 모두 절편과 분산의 사전분포에 대한 언급은 없었다. 이에 먼저 절편 모수에 대해서는 모수가 존재할만한 범위에 균등분포를 한정하는 방법을 사용하였다. 측정도구의 반응 개수를 기준으로 매개변수에는 $U[1, 7]$, 종속변수에는 $U[1, 5]$ 를 사전분포로 할당하였다.

다음으로 분산 모수 역시 추정치를 보고하는 연구를 찾을 수 없어 434명의 대학생을 대상으로 사회부과 완벽주의가 사회불안에 미치는 영향을 대인존재감과 자기수용이 매개하는 모형을 연구한 Choi와 Hong(2020)의 확인적 요인분석 결과를 활용하였다. 보고된 요인 간 상관과 요인부하를 이용하면 모형함의 공분산 행렬(model-implied covariance matrix)을 계산할

수 있으므로 이를 통해 확률적으로 100만 개의 자료를 생성하였다. 생성한 데이터를 이용해 예시와 같은 매개모형을 추정하였다. 분산 추정치인 s_1^2 과 s_2^2 은 각각 0.555와 0.524로 확인되었다. 이 추정치를 식 7에 대입해 역감마분포의 하이퍼파라미터를 계산하였다. 무조건적 자기수용의 분산 모수 하이퍼파라미터는 $\alpha = 217, \beta = 120.435$, 사회불안의 분산 모수 하이퍼파라미터는 $\alpha = 217, \beta = 113.708$ 로 계산되었다. 그러나 이 값으로 역감마분포를 지정하면 폭이 매우 좁아 정보 사전분포로 기능할 수 있으므로 본래의 의도를 만족시키기 어렵다. 따라서 본 예시에서는 선행연구의 표본크기를 1/10로 가정하고 분포의 불확실성을 높이고자 하였다. 결과적으로 무조건적 자기수용의 잔차분산 모수에는 $\Gamma^{-1}(21.7, 12.4)$, 사회불안의 잔차분산 모수에는 $\Gamma^{-1}(21.7, 11.7)$ 이 할당되었다.

정보 사전분포 설정

각 경로계수에 대한 정보 사전분포 설정을 위해 본 연구의 예시 변수와 비슷한 변수를 사용한 연구를 참고하였다. 선행연구들은 연구마다 다른 척도를 사용하였으므로, 표준화 추정치 결과를 이용하였다. Seo et al. (2018)은 대학생을 대상으로 평가염려 완벽주의가 사회불안에 미치는 영향을 무조건적 자기수용이 매개하는 모형을 연구하였다. 매개모형의 추정치는 경로 $a(\hat{\beta} = -0.27, p < .001)$, 경로 $b(\hat{\beta} = -0.16, p < .001)$, 경로 $c'(\hat{\beta} = 0.52, p < .001)$ 이었고, 이는 그림 10과 같다.

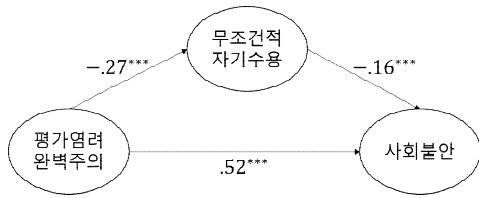


그림 10. Seo et al. (2018)의 모형 결과

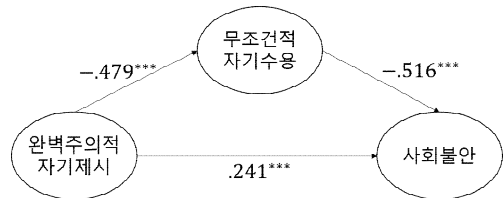


그림 12. Lee(2020)의 모형 결과

Choi와 Hong(2020)이 대학생을 대상으로 연구한 매개모형의 추정치는 경로 a ($\hat{\beta} = -0.487, p < .001$), 경로 b ($\hat{\beta} = -0.329, p < .001$), 경로 c' ($\hat{\beta} = -0.015, p = .081$)이었고, 이는 그림 11과 같다. 다만, Choi와 Hong(2020)의 연구 모형은 사회부과 완벽주의가 사회불안에 미치는 영향을 대인존재감과 무조건적 자기수용이 매개하는 이중매개모형이며, 본 연구는 결과 중 필요한 부분을 발췌한 것이므로 단순매개모형을 상정하였을 때의 결과와 다소 차이가 있을 수 있음을 고려해야만 하였다.

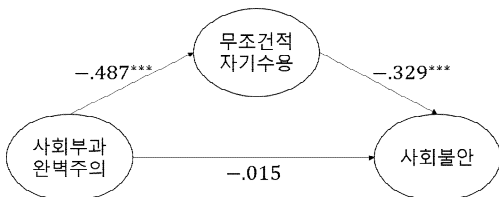


그림 11. Choi와 Hong(2020)의 모형 결과

청소년을 대상으로 자신이 타인에게 어떻게 드러나는가를 나타내는 완벽주의적 자기제시가 사회불안에 미치는 영향을 무조건적 자기수용이 매개하는 모형을 연구한 Lee(2020)의 추정치는 경로 a ($\hat{\beta} = -0.479, p < .001$), 경로 b ($\hat{\beta} = -0.516, p < .001$), 경로 c' ($\hat{\beta} = 0.241, p < .001$)로 그림 12와 같다.

사회부과 완벽주의가 우울에 미치는 영향을 무조건적 자기수용이 매개하는 모형을 연구한 Choi와 Song(2018)의 연구도 사전분포 설정에 참고하였는데, 이는 불안과 우울의 상관성이 높기 때문이다(Steer & Clark, 1997). Choi와 Song(2018)의 연구 결과에서는 경로 a ($\hat{\beta} = -0.95, p < .001$), 경로 b ($\hat{\beta} = -0.45, p < .001$)의 추정치만 유의하였고, 그림 13과 같다.



그림 13. Choi와 Song(2018)의 모형 결과

먼저, 간접효과와 관련된 사전분포의 평균 하이퍼파라미터는 네 개의 선행연구 추정치 평균으로 설정하였다. 각 경로 추정치의 평균은 $a = -0.547, b = -0.364$ 이다. 직접효과의 사전분포는 직접효과가 유의하게 확인된 Seo et al. (2018)과 Lee(2020)의 연구를 참고하였다. Lee(2020)의 연구는 청소년이 대상이었으므로 Seo et al. (2018)의 추정치와 더 가까운 0.4를 평균 하이퍼파라미터로 선택하였다. 다음으로 강한 정보 사전분포의 분산 하이퍼파라미터는 앞서 제시한 경향에 따라 평균 하이

퍼파라미터 값의 10%로 설정하였다. 따라서 경로 a, b, c' 의 강한 정보 사전분포는 각각 $N(-0.55, 0.055)$, $N(-0.36, 0.036)$, $N(0.4, 0.04)$ 이었다. 다음으로, 약한 정보 사전분포는 경향에 따라 평균 하이퍼파라미터의 100%보다는 작으면서도 강한 정보 사전분포와는 차이를 줄 수 있는 값을 선택하고자 평균 하이퍼파라미터의 50%로 설정하였다. 따라서 지정된 약한 정보 사전분포는 $N(-0.55, 0.275)$, $N(-0.36, 0.18)$, $N(0.4, 0.2)$ 이었다. 각 사전분포는 표 3에 제공된다.

이렇게 결정한 값들을 적용하여 각 모수의 사전분포를 설정하였다. 메타 분석, 리뷰 논문, 전문가 조언 등의 정보는 얻지 못하였으나, 표본과 모형이 비슷한 몇 개의 경험적 연구만으로도 직접 사전분포를 설정하는 것이 가능함을 보이고자 하였다. ML 방법 및 디폴트, 약한 정보, 강한 정보 사전분포를 지정해 베이지안 방법을 사용하여 얻은 분석 결과는 표 4에 제시되어 있다. 결과 추정치에서 두드러진 차이는 발견되지 않았다. 이는 복잡하지 않은 모형 추정에 378이라는 작지 않은 표본

크기로 인해 모든 추정이 비슷하게 수렴했다고 해석할 수 있다. 다만, 베이지안 추정 결과 ppp(posterior predictive p-values) 모형적합도에서 사전분포를 좁은 범위로 한정된 강한 정보 사전분포가 가장 좋지 않은 모형적합도를 보였으며, 사전 정보를 활용하여 약한 정보 사전분포를 설정한 결과가 가장 좋은 모형적합도를 보였다. ppp는 베이지안 모형적합도 확인의 한 방법이며, 모형이 자료에 적합할수록 0.5에 가까운 값이 나타난다(Gelman et al., 2013).

결론 및 논의

최근 베이지안 추정에 대한 관심이 높아짐에 따라 베이지안 방법을 이용한 연구가 계속 증가하는 추세에 있으나(van de Schoot et al., 2017), 정작 베이지안 방법에서 가장 중요한 사전분포의 지정은 여전히 간과되고 있다. 이는 사전분포의 설정에 큰 신경을 기울이지 않고 프로그램이 제공하는 디폴트 사전분포만을

표 3. 예시에 사용된 사전분포

모수 (Parameter)	사전분포(Priors)			
	무정보 (Noninformative)	약한 정보 (Weakly informative)	강한 정보 (Strong informative)	디폴트 (Default)
절편 i_1	$U[1, 7]$	-	-	$N(0, 10^{10})$
절편 i_2	$U[1, 5]$	-	-	$N(0, 10^{10})$
경로 a	-	$N(-.55, .275)$	$N(-.55, .055)$	$N(0, 10^{10})$
경로 b	-	$N(-.36, .18)$	$N(-.36, .036)$	$N(0, 10^{10})$
경로 c'	-	$N(.4, .2)$	$N(.4, .04)$	$N(0, 10^{10})$
분산 $\sigma_{d_1}^2$	$\Gamma^{-1}(21.7, 12.04)$	-	-	$\Gamma^{-1}(-1, 0)$
분산 $\sigma_{d_2}^2$	$\Gamma^{-1}(21.7, 11.37)$	-	-	$\Gamma^{-1}(-1, 0)$

표 4. 추정 결과

모수	빈도주의		베이지안	
	최대우도방법	약한 정보 사전분포	강한 정보 사전분포	디폴트 사전분포
절편 i_1	1.027* (1.332)	1.028* (1.327)	1.028* (1.329)	1.026* (1.325)
절편 i_2	7.725*** (8.953)	7.713*** (8.929)	7.665*** (8.946)	7.724*** (8.925)
경로 a	-1.041*** (-0.751)	-1.037*** (-0.747)	-1.022*** (-0.742)	-1.041*** (-0.748)
경로 b	-0.247*** (-0.276)	-0.247*** (-0.275)	-0.247*** (-0.273)	-0.247*** (-0.275)
경로 c'	0.493*** (0.398)	0.493*** (0.396)	0.493*** (0.396)	0.493*** (0.396)
분산 $\sigma_{d_1}^2$	0.357*** (0.601)	0.363*** (0.604)	0.363*** (0.606)	0.363*** (0.603)
분산 $\sigma_{d_2}^2$	0.325*** (0.437)	0.329*** (0.442)	0.329*** (0.449)	0.329*** (0.440)
ppp	완전포화모형	0.499	0.487	0.496

주. 추정치는 비표준화 추정치(B)와 표준화 추정치(β)이며, 괄호 안의 값이 표준화 추정치임. ppp = posterior predictive p-values. * $p < .05$; *** $p < .001$.

이용한 연구들이나, 어떤 사전분포를 지정하였는지 밝히지 않는 연구들이 존재하는 것을 통해 짐작해볼 수 있는 문제였다. 따라서 본 연구는 사전분포 및 사전분포의 지정과 관련된 개념을 종합하여 정리하고, 이후 직접 지정하지 않은 디폴트 사전분포와 직접 지정한 무정보 및 정보 사전분포를 나누어 베이지안 추정의 방법과 절차를 전개함으로써 사전분포와 관련된 쟁점에 대해 논의하는 것을 목적으로 하였다. 구체적으로는 사전분포 지정이 베이지안 추정에서 가장 깊은 주의를 요하는 과정 중 하나임을 지적하고, 나아가 사전분포를 직접 지정할 때 실질적으로 참고할 수 있는

방법을 제공하고자 하였다. 이러한 목적을 달성하기 위해 디폴트 사전분포에 대해 탐색했을뿐만 아니라, 이의 대안으로 무정보 사전분포의 지정에 참고할 수 있는 방법 및 사전분포의 정보성 경향에 대해 논의하였다. 또한 사전분포를 직접 지정하는 예시를 통해 실제 연구에서의 활용에도 도움을 주고자 하였다.

앞서 다룬 본문의 논의를 종합해 정리해보면, 첫째로 디폴트 사전분포는 여러 문제점을 수반할 위험이 있으므로 베이지안 방법을 사용하는 연구자는 모수에 대한 정보가 없을 때, 디폴트 사전분포가 갖는 특징에 대해 명확히 이해한 후, 디폴트 사전분포를 사용할지

무정보 사전분포를 사용할지 고민하는 과정이 필요하다. 이때 무정보 사전분포를 설정하고자 한다면 참고할 수 있는 몇 가지 방법이 존재하며, 이는 본문에 자세히 정리되어 있다. 무정보 사전분포의 설정 방식은 한 가지로 정해져 있는 것이 아니므로 연구자의 관점에 따라 여러 방향에서 접근이 가능하지만, 본 연구는 참조할 수 있는 몇 개의 지침을 제공함으로써 연구자들이 무분별하게 디폴트 사전분포를 사용하는 것을 지양하도록 하는 데 소정의 목적을 두었다. 둘째, 현재까지 연구된 여러 문헌의 직접 지정된 사전분포를 취합해 사전분포 정보성의 경향을 찾아 가이드라인을 제공함으로써 위치 모수의 사전분포 지정에 도움을 줄 수 있는 정보를 제공하였다. 이 결과는 여러 시뮬레이션 연구와 경험적 연구를 통합하여 다루었다는 점에서 참고할 만한 가치가 있다. 현재 수행되는 연구마다 합의되지 않은 기준으로 사전분포의 정보성을 정하는 기조는 기존 연구를 참고해 베이지안 추정을 사용하고자 하는 연구자들에게 적절한 지침이 되지 못한다. 이에 최근 연구자들의 경향을 파악할 수 있는 본 연구는 사전분포의 설정에 실질적인 도움이 될 수 있을 것이다. 마지막으로 실제 자료를 바탕으로 찾은 사전분포의 경향성 기준을 적용한 실제 예시를 함께 제공하였다. 이는 간단한 예를 통해, 직접적이고 구체적인 정보가 없더라도 적절히 기능하는 정보 사전분포를 지정할 수 있음을 보이고자 한 것이었다. 또한 ML 방법과 베이지안 방법의 결과 비교를 통해 제한된 사전 정보만으로도 두 추정이 비슷한 결과를 산출한다는 것을 확인하였다. 더불어 본문에서 언급했던 정확하지 않은 정보로 설정한 사전분포라면 약한 정보 사전분포를 설정하는 것이 더 나은 배

이지안 방법의 모형적합도인 ppp 값을 통해 확인하였다.

본 연구를 통해 사전분포의 중요성과 사전분포 지정에 필요한 전반적인 지식을 전달하였지만, 몇 가지 제한점이 있다. 먼저, 본문에서는 사전분포의 경향을 조사할 때, 정규분포를 이용한 위치 모수의 사전분포 기준만을 논의하였다. 본 연구의 주요 목적이 베이지안 연구를 진행하고자 하는 연구자들의 사전분포 설정을 돕는 것에 있었으므로 일반적으로 가장 많이 이용되는 관심 모수에 관한 연구가 선행되어야 한다고 판단하였다. 또한 분산 모수에 대한 사전분포나 정규분포 이외의 사전분포는 정보성 경향을 파악하기에 아직은 학계에 쌓인 자료가 부족하다는 한계가 있었다. 향후 사전분포 지정에 대한 연구가 더욱 활발해지고, 위치 모수에 다른 종류의 분포를 이용해 사전분포를 지정한 연구나 분산 모수에 사전분포를 직접 지정한 연구가 늘어난다면, 이에 대한 경향성을 다루는 것 역시 가능할 것이다. 다음으로 본 연구에서 예시에 사용한 프로그램인 *Mplw*는 공액 분포만 사전분포로 지정할 수 있는 제한이 있다. 이에 본문에서 다룬 Half-Cauchy와 같이 공액을 갖지 않는 분포는 예시에서 사용할 수 없었다. 다만 본 연구는 여러 분포를 지정해보고 그 수행도(performance)를 비교하는 데 목적이 있는 시뮬레이션 연구가 아니었으므로 추후 분포의 종류에 따라 수행 정도를 비교하는 등의 연구를 위해서는 다른 프로그램을 통한 접근이 필요할 것으로 판단된다.

앞서 언급한 제한점에도 불구하고, 본 연구는 구조방정식 분야에서 상대적으로 새로운 추정 방법인 베이지안 방법에 대한 이해를 높이고, 사전분포 지정의 중요성을 강조하기 위

해 사전분포의 기본적이고 구체적인 개념을 종합적으로 다루었으며, 정보 사전분포의 하이퍼파라미터가 설정되는 경향을 파악하고자 한 첫 번째 연구라는데 그 의의가 있다. 이를 통해 연구자들이 베이지안 방법을 사용하고자 할 때, 사전분포를 이용하기 위해 분석 전 고려해야 할 점이 무엇인지 파악할 수 있을 것이며, 사전분포를 직접 지정하는 과정도 더욱 수월하고 체계적으로 이루어질 것을 기대한다.

참고문헌

- Asparouhov, T., & Muthén, B. (2010). Bayesian analysis of latent variable models using Mplus. Retrieved from <http://www.statmodel.com/download/BayesAdvantages18.pdf>
- Baldwin, S. A., & Fellingham, G. W. (2013). Bayesian methods for the analysis of small sample multilevel data with a complex variance structure. *Psychological Methods, 18*(2), 151-164. <https://doi.org/10.1037/a0030642>
- Barnard, J., McCulloch, R., & Meng, X. L. (2000). Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage. *Statistica Sinica, 10*(4), 1281-1311. Retrieved from <http://www3.stat.sinica.edu.tw/statistica/oldpdf/a10n416.pdf>
- Baron, R. M., & Kenny, D. A. (1986). The moderator - mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Browne, W. J., & Draper, D. (2006). A comparison of Bayesian and likelihood-based methods for fitting multilevel models. *Bayesian Analysis, 1*(3), 473-514. <https://doi.org/10.1214/06-BA117>
- Can, S., van de Schoot, R., & Hox, J. (2015). Collinear latent variables in multilevel confirmatory factor analysis: A comparison of maximum likelihood and Bayesian estimations. *Educational and Psychological Measurement, 75*(3), 406-427. <https://doi.org/10.1177/0013164414547959>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... & Riddell, A. (2017). Stan: A probabilistic programming language. *Grantee Submission, 76*(1), 1-32. <https://doi.org/10.18637/jss.v076.i01>
- Chiorri, C., Day, T., & Malmberg, L. E. (2014). An approximate measurement invariance approach to within-couple relationship quality. *Frontiers in Psychology, 5*, 983. <https://doi.org/10.3389/fpsyg.2014.00983>
- Choi, J. G. & Song, W. Y. (2018). The Effect of Socially-Prescribed Perfectionism of College Students to Depression: Testing the Mediation effect of Intolerance of Uncertainty and Unconditional Self-Acceptance. *Journal of Convergence for Information Technology, 8*(3), 183-191. <https://doi.org/10.22156/CS4SMB.2018.8.3.183>
- Choi, Y. J. & Hong, H. Y. (2020). Mediating Effects of Mattering and Self-Acceptance in

- the Relationship between Socially Prescribed Perfectionism and Social Anxiety. *Journal of the Korea Contents* 20(1), 259-270.
<https://doi.org/10.5392/JKCA.2020.20.01.259>
- Cortopassi, A. C., Starks, T. J., Parsons, J. T., & Wells, B. E. (2017). Self-concealment, ego depletion, and drug dependence among young sexual minority men who use substances. *Psychology of Sexual Orientation and Gender Diversity*, 4(3), 272-281.
<https://doi.org/10.1037/sgd0000230>
- Depaoli, S. (2012). Measurement and structural model class separation in mixture CFA: ML/EM versus MCMC. *Structural Equation Modeling: A Multidisciplinary Journal*, 19(2), 178-203.
<https://doi.org/10.1080/10705511.2012.659614>
- Depaoli, S. (2014). The impact of inaccurate “informative” priors for growth parameters in Bayesian growth mixture modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(2), 239-252.
<https://doi.org/10.1080/10705511.2014.882686>
- Depaoli, S., & Clifton, J. P. (2015). A Bayesian approach to multilevel structural equation modeling with continuous and dichotomous outcomes. *Structural Equation Modeling: A Multidisciplinary Journal*, 22(3), 327-351.
<https://doi.org/10.1080/10705511.2014.937849>
- Falkenström, F., Hatcher, R. L., & Holmqvist, R. (2015a). Confirmatory Factor Analysis of the Patient Version of the Working Alliance Inventory - Short Form Revised. *Assessment*, 22(5), 581 - 593.
<https://doi.org/10.1177/1073191114552472>
- Falkenström, F., Hatcher, R. L., Skjulsvik, T., Larsson, M. H., & Holmqvist, R. (2015b). Development and validation of a 6-item working alliance questionnaire for repeated administrations during psychotherapy. *Psychological Assessment*, 27(1), 169 - 183.
<https://doi.org/10.1037/pas0000038>
- Fang, J., Wen, Z., & Hau, K. T. (2019). Mediation effects in 2-1-1 multilevel model: evaluation of alternative estimation methods. *Structural Equation Modeling: A Multidisciplinary Journal*, 26(4), 591-606.
<https://doi.org/10.1080/10705511.2018.1547967>
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*, 1(3), 515-534.
<https://doi.org/10.1214/06-BA117A>
- Gelman, A., Carlin, J. B., & Stern, H. S. (2013). *Bayesian Data Analysis* (3rd ed.). CRC Press.
<https://doi.org/10.1201/9780429258411>
- Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models. *Annals of Applied Statistics*, 2(4), 1360-1383.
<https://doi.org/10.1214/08-AOAS191>
- Hagger, M. S., & Hamilton, K. (2018). Motivational predictors of students’ participation in out-of-school learning activities and academic attainment in science: An application of the trans-contextual model using Bayesian path analysis. *Learning and Individual Differences*, 67, 232-244.
<https://doi.org/10.1016/j.lindif.2018.09.002>

- Hobert, J. P., & Casella, G. (1996). The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. *Journal of the American Statistical Association*, 91(436), 1461-1473.
<https://doi.org/10.1080/01621459.1996.10476714>
- Holtmann, J., Koch, T., Lochner, K., & Eid, M. (2016). A comparison of ML, WLSMV, and Bayesian methods for multilevel structural equation models in small samples: A simulation study. *Multivariate Behavioral Research*, 51(5), 661-680.
<https://doi.org/10.1080/00273171.2016.1208074>
- Hox, J. J., & Bechger, T. M. (1998). An introduction to structural equation modeling. *Family Science Review*, 11, 354-373. Retrieved from <http://joophox.net/publist/semfamre.pdf>
- Hox, J. (2020). Important yet unheeded. In van de Schoot, R., & Miočević, M. (Eds.). *Small sample size solutions: A guide for applied researchers and practitioners* (pp. 255-265). Taylor & Francis. Retrieved from <http://library.oapen.org/handle/20.500.12657/22385>
- Hoyle, R. H. (2000). Confirmatory Factor Analysis. In Tinsley, H. E., & Brown, S. D. (Eds.). *Handbook of Applied Multivariate Statistics and Mathematical Modeling* (pp. 465-497). Academic Press.
<https://doi.org/10.1016/B978-0-12-691360-6.X5000-9>
- Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. Proceedings of the Royal Society of London. Series A. *Mathematical and Physical Sciences*, 186(1007), 453-461.
<https://doi.org/10.1098/rspa.1946.0056>
- Kaplan, D. & Depaoli, S. (2013). Bayesian statistical methods. In T. D. Little (Ed.), *Oxford handbook of quantitative methods* (pp. 407-437). Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199934874.001.0001>
- Kelly, B. C., Rendina, H. J., Vuolo, M., Wells, B. E., & Parsons, J. T. (2015a). Influences of motivational contexts on prescription drug misuse and related drug problems. *Journal of Substance Abuse Treatment*, 48, 49-55.
<https://doi.org/10.1016/j.jsat.2014.07.005>
- Kelly, B. C., Vuolo, M., Pawson, M., Wells, B. E., & Parsons, J. T. (2015b). Chasing the bean: Prescription drug smoking among socially active youth. *The Journal of Adolescent Health*, 56, 632-638.
<https://doi.org/10.1016/j.jadohealth.2015.02.008>
- Kim, S. Y., Suh, Y., Kim, J. S., Albanese, M. A., & Langer, M. M. (2013). Single and multiple ability estimation in the SEM framework: A noninformative Bayesian estimation approach. *Multivariate Behavioral Research*, 48, 563-591.
<https://doi.org/10.1080/00273171.2013.802647>
- Kim, S. Y., Huh, D., Zhou, Z., & Mun, E. Y. (2020). A comparison of Bayesian to maximum likelihood estimation for latent growth models in the presence of a binary outcome. *International Journal of Behavioral Development*, 44(5), 447-457.
<https://doi.org/10.1177/0165025419894730>
- Kruschke, J. K. (2011). Introduction to special section on Bayesian data analysis. *Perspectives*

- on *Psychological Science*, 6(3), 272-273.
<https://doi.org/10.1177/1745691611406926>
- Lee, D. Y. (2020). *The Effects of adolescent's Perfectionistic Self-Presentation on Social anxiety: focusing on the mediation effects of unconditional self-acceptance* [Unpublished master's thesis]. The Catholic University of Korea.
- Lee, J. E. & Kim, S. Y. (2018). The Mediating Effects of Unconditional Self-Acceptance and Moderating Effects of Gender on the Relationship between Evaluative Concerns Perfectionism and Social Anxiety among College Students. *Journal of Human Understanding and Counseling*, 39(2), 25-45.
<https://doi.org/10.30593/JHUC.39.2.2>
- Lee, S. Y., & Song, X. Y. (2004). Evaluation of the Bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes. *Multivariate Behavioral Research*, 39(4), 653-686.
https://doi.org/10.1207/s15327906mbr3904_4
- Lüdtke, O., Marsh, H. W., Robitzsch, A., & Trautwein, U. (2011). A 2×2 taxonomy of multilevel latent contextual models: Accuracy - bias trade-offs in full and partial error correction models. *Psychological Methods*, 16(4), 444-467. <https://doi.org/10.1037/a0024376>
- Lunn, D., Jackson, C., Best, N., Thomas, A., & Spiegelhalter, D. (2012). *The BUGS book: A practical introduction to Bayesian analysis*. CRC Press.
<https://doi.org/10.1080/09332480.2013.868759>
- Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, 10(4), 325-337.
<https://doi.org/10.1023/A:1008929526011>
- McNeish, D. (2016). On using Bayesian methods to address small sample problems. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(5), 750-773.
<https://doi.org/10.1080/10705511.2016.1186549>
- McNeish, D. (2019) Two-Level Dynamic Structural Equation Models with Small Samples. *Structural Equation Modeling: A Multidisciplinary Journal*, 26(6), 948-966.
<https://doi.org/10.1080/10705511.2019.1578657>
- McNeish, D., & Stapleton, L. M. (2016). Modeling clustered data with very few clusters. *Multivariate Behavioral Research*, 51(4), 495-518.
<https://doi.org/10.1080/00273171.2016.1167008>
- Merkle, E. C., & Rosseel, Y. (2015). blavaan: Bayesian structural equation models via parameter expansion. *Journal of Statistical Software*, 85(4), 1-30.
<https://doi.org/10.18637/jss.v085.i04>
- Miočević, M., Levy, R., & MacKinnon, D. P. (2021) Different Roles of Prior Distributions in the Single Mediator Model with Latent Variables, *Multivariate Behavioral Research*, 56(1), 20-40,
<https://doi.org/10.1080/00273171.2019.1709405>
- Miočević, M., MacKinnon, D. P., & Levy, R. (2017). Power in Bayesian mediation analysis for small sample research. *Structural equation modeling: a multidisciplinary journal*, 24(5), 666-683.
<https://doi.org/10.1080/10705511.2017.1312407>

- Mohajerin, B., Dolatshahi, B., Pour Shahbaz, A., & Farhoudian, A. (2013). Differences between expressive suppression and cognitive reappraisal in opioids and stimulant dependent patients. *International Journal of High Risk Behaviors & Addiction, 2*, 8 - 14. <https://doi.org/10.5812/ijhrba.8514>
- Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: a more flexible representation of substantive theory. *Psychological Methods, 17*(3), 313-335. <https://doi.org/10.1037/a0026802>
- Muthen, L. K., & Muthen, B. O. (1998 - 2019). *Mplus user's guide*. (8th ed.) [Computer software manual]. Muthen & Muthen. Retrieved from https://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf
- Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In K. Hornik, F. Leisch, & A. Zeileis (Eds.), In *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124). Technische Universität Wien. Retrieved from <https://www.r-project.org/conferences/DSC-2003/Proceedings/>
- Polson, N. G., & Scott, J. G. (2012). On the half-Cauchy prior for a global scale parameter. *Bayesian Analysis, 7*(4), 887-902. <https://doi.org/10.1214/12-BA730>
- Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42*(1), 185-227. <https://doi.org/10.1080/00273170701341316>
- R Core Team. (2015). *R: A language and environment for statistical computing* [Computer software manual]. R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Seo, Y. I., Kim, J. M., & Jo, H. B. (2018). The effects of multidimensional perfectionism on the social anxiety and depression of college students: The mediating effect of unconditional self-acceptance. *Korea Society for the Emotional & Behavioral Disorders, 34*(1), 197-215. <https://doi.org/10.33770/JEBD.34.1.16>
- Smid, S. C., McNeish, D., Miočević, M., & van de Schoot, R. (2020a). Bayesian Versus Frequentist estimation for Structural Equation Models in Small Sample Contexts: A Systematic Review. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(1), 131-161. <https://doi.org/10.1080/10705511.2019.1577140>
- Smid, S. C., Depaoli, S., & van De Schoot, R. (2020b). Predicting a distal outcome variable from a latent growth model: ML versus Bayesian estimation. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(2), 169-191. <https://doi.org/10.1080/10705511.2019.1604140>
- Smid, S. C., & Winter, S. D. (2020). Dangers of the Defaults: A Tutorial on the Impact of Default Priors When Using Bayesian SEM With Small Samples. *Frontiers in Psychology, 11*, 3536. <https://doi.org/10.3389/fpsyg.2020.611963>

- Spiegelhalter, D. J., Thomas, A., Best, N. G., & Lunn, D. (2003). *WinBUGS user manual, Version 1.4*. MRC Biostatistics Unit. Retrieved from <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/manual14.pdf>
- Steer, R. A., & Clark, D. A. (1997). Psychometric characteristics of the Beck Depression Inventory-II with college students. *Measurement and Evaluation in Counseling and Development*, 30(3), 128-136. <https://doi.org/10.1080/07481756.1997.12068933>
- van de Schoot, R., Broere, J. J., Perryck, K. H., Zondervan-Zwijnenburg, M., & Loey, N. E. (2015). Analyzing small data sets using Bayesian estimation: The case of posttraumatic stress symptoms following mechanical ventilation in burn survivors. *European Journal of Psychotraumatology*, 6(1), Article 25216. <https://doi.org/10.3402/ejpt.v6.25216>
- van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & Aken, M. A. (2014). A gentle introduction to Bayesian analysis: Applications to developmental research. *Child development*, 85(3), 842-860. <https://doi.org/10.1111/cdev.12169>
- van de Schoot, R., Winter, S. D., Ryan, O., Zondervan-Zwijnenburg, M., & Depaoli, S. (2017). A systematic review of Bayesian articles in psychology: The last 25 years. *Psychological Methods*, 22(2), 217-239. <https://doi.org/10.1037/met0000100>
- van Erp, S., Mulder, J., & Oberski, D. L. (2018). Prior sensitivity analysis in default Bayesian structural equation modeling. *Psychological Methods*, 23(2), 363-388. <https://doi.org/10.1037/met0000162>
- Xu, C., Wang, W., Liu, P., & Li, Z. (2015). Calibration of crash risk models on freeways with limited real-time traffic data using Bayesian meta-analysis and Bayesian inference approach. *Accident Analysis & Prevention*, 85, 207-218. <https://doi.org/10.1016/j.aap.2015.09.016>
- Yu, R., & Abdel-Aty, M. (2013). Investigating different approaches to develop informative priors in hierarchical Bayesian safety performance functions. *Accident Analysis & Prevention*, 56, 51-58. <https://doi.org/10.1016/j.aap.2013.03.023>
- Yuan, Y., & MacKinnon, D. P. (2009). Bayesian mediation analysis. *Psychological Methods*, 14(4), 301-322. <https://doi.org/10.1037/a0016972>
- Zondervan-Zwijnenburg, M., Peeters, M., Depaoli, S., & van de Schoot, R. (2017). Where do priors come from? Applying guidelines to construct informative priors in small sample research. *Research in Human Development*, 14(4), 305-320. <https://doi.org/10.1080/15427609.2017.1370966>

1차원고접수 : 2021. 08. 17.

2차원고접수 : 2021. 10. 25.

최종게재결정 : 2021. 11. 06.

Understanding and applying prior distributions in Bayesian analyses

Ji-Yoon Lee

Su-Young Kim

Department of Psychology, Ewha Womans University

The Bayesian estimation method has recently received a lot of attention in the social sciences. The Bayesian method has a special factor of prior distribution that can reflect researchers' background knowledge in the estimation process. The specification of the prior distribution affects the overall estimation. Despite prior distribution being the most important factor in Bayesian analysis, there is a lack of methodological research for understanding and appropriately specifying the prior distribution. Therefore, the present study tries to help researchers to apply the prior distribution to their estimation by addressing the importance of the prior distribution and the overall content of the prior specification. First, we explore the method that researchers do not directly specify the prior distribution. This method means selecting the default prior distribution automatically provided by the program, and if you want to use this option, you must know exactly what kind of the default prior distribution is actually provided. For this, we discuss the default priors of frequently used programs, as well as the known problem of the default priors. Second, we address the method that researchers do specify the prior distribution by themselves. The prior distributions that can be directly specified include noninformative prior distributions and informative prior distributions. Which prior distribution to use is determined by the presence of prior information on parameters. This study deals with the necessity of noninformative prior distributions and the proposed method when specifying them, provides studies that can be referenced when specifying informative prior distributions, and explores criteria that can be referenced for the select of informativeness by synthesizing the criteria across many studies. We provide practical help through data examples applying the methods discussed in the text, and finally discuss the significance and limitations of the present study.

Keywords: Bayesian method, prior distribution, default priors, non-informative priors, informative priors