



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0143158
(43) 공개일자 2016년12월14일

(51) 국제특허분류(Int. Cl.)
G06F 9/455 (2006.01) G06F 17/00 (2006.01)
(52) CPC특허분류
G06F 9/45504 (2013.01)
G06F 17/00 (2013.01)
(21) 출원번호 10-2015-0079350
(22) 출원일자 2015년06월04일
심사청구일자 2015년06월04일

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
최태런
서울특별시 성북구 종암로23길 35, 205동 602호 (종암동, 래미안세레니티)
조성일
서울특별시 동작구 양녕로20길 127, 404호 (상도동, 한신베네스트)
김형준
서울특별시 서초구 명달로1길 51, 102동 902호 (방배동, 방배어울림)
(74) 대리인
특허법인엠에이피에스

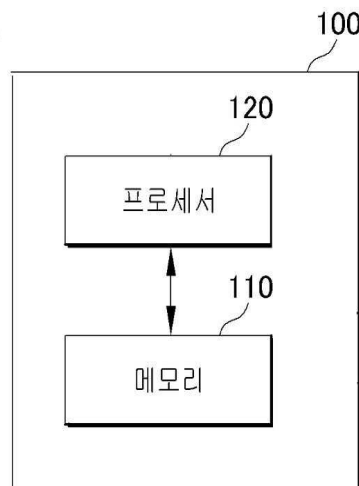
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 용량 반응곡선의 추정 및 기준용량 계산을 위한 장치 및 방법

(57) 요약

본 발명은 용량반응곡선의 추정 및 기준용량을 계산하는 장치 및 방법에 대하여 개시하고 있다. 본 발명의 일 실시예에 따른 용량 반응곡선 추정 및 기준용량 계산 장치는 사용자의 입력값에 따라 용량반응곡선을 추정하고, 기준용량을 계산하는 프로그램을 저장하는 메모리; 및 프로그램을 실행하는 프로세서를 포함하되, 프로세서는 프로그램의 수행에 따라, 분석대상에 대하여 사전 분포를 설정하고, 용량반응곡선에 대하여 형태제약을 부과하고, 부과된 형태제약에 따른 용량반응모형에 대하여 베이지안 방법론에 따라 사후분포를 계산하고, 계산된 사후분포로부터 임의표본을 추출하고, 추출된 임의표본을 통해 용량반응곡선의 추정하고, 추정된 용량반응곡선에 대한 기준용량을 계산한다. 이때, 분석대상은 비모수 함수 회귀모형으로 가정한다.

대표도 - 도1



(52) CPC특허분류

G06F 17/10 (2013.01)

G06F 17/18 (2013.01)

G06N 7/005 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1345225334

부처명 교육부

연구관리전문기관 한국연구재단

연구사업명 일반연구자지원

연구과제명 대규모 자료를 위한 준모수적 베이지안 회귀분석에 관한 연구

기 여 율 9/10

주관기관 고려대학교

연구기간 2013.09.01 ~ 2015.08.31이 발명을 지원한 국가연구개발사업

과제고유번호 1345210124

부처명 미래창조과학부

연구관리전문기관 한국연구재단

연구사업명 일반연구자지원(미래부)

연구과제명 비모수적 적합도 검정에서의 통합적 통계 방법론 연구

기 여 율 1/10

주관기관 고려대학교

연구기간 2013.12.01 ~ 2016.11.30

명세서

청구범위

청구항 1

용량반응곡선의 추정 및 기준용량을 계산하는 장치에 있어서,

사용자의 입력값에 따라 상기 용량반응곡선을 추정하고, 상기 기준용량을 계산하는 프로그램을 저장하는 메모리; 및

상기 프로그램을 실행하는 프로세서를 포함하되,

상기 프로세서는 상기 프로그램의 수행에 따라, 분석대상에 대하여 사전 분포를 설정하고, 용량반응곡선에 대한 형태제약을 부과하고, 부과된 형태제약에 따른 용량반응모형에 대하여 베이지안 방법론에 따라 사후분포를 계산하고, 상기 계산된 사후분포로부터 임의표본을 추출하고, 상기 추출된 임의표본을 통해 용량반응곡선의 추정하고, 추정된 상기 용량반응곡선에 대한 상기 기준용량을 계산하며,

상기 분석대상은 비모수 함수 회귀모형으로 가정하는 것인, 용량반응곡선의 추정 및 기준용량을 계산하는 장치.

청구항 2

제 1항에 있어서,

상기 프로세서는

상기 프로그램에서 제공하는 사용자 인터페이스를 통해 상기 사용자로부터 분석대상, 형태제약, 기준용량계산방법에 대한 입력값을 입력받는 것인, 용량반응곡선의 추정 및 기준용량을 계산하는 장치.

청구항 3

제 2항에 있어서,

상기 프로세서는

상기 사용자 인터페이스를 통해 사용자로부터 단조증가, 단조감소, 오목, 볼록 및 s자의 형태제약 중 하나를 입력받는 것인, 용량반응곡선의 추정 및 기준용량을 계산하는 장치.

청구항 4

제 1항에 있어서,

상기 프로세서는

상기 프로그램에서 제공하는 출력 인터페이스를 통해 상기 프로그램 수행에 따른 상기 용량반응곡선 및 상기 기준용량계산에 대한 결과를 제공하는 것인, 용량 반응 곡선의 추정 및 기준용량을 계산하는 장치.

청구항 5

제 4항에 있어서,

상기 프로세서는 상기 출력인터페이스를 통해 상기 용량반응곡선 및 상기 기준용량 계산 결과에 대한 요약 통계를 제공하는 것인, 용량반응곡선의 추정 및 기준용량을 계산하는 장치.

청구항 6

제 4항에 있어서,

상기 프로세서는 상기 출력 인터페이스를 통해 트레이스 플롯을 제공하고, 상기 플롯 트레이스는 상기 용량반응곡선 및 상기 기준용량계산에 대한 결과에 있어 비모수들의 수렴여부를 확인하기 위해 제공되는 것인, 용량반응곡선의 추정 및 기준용량을 계산하는 장치.

청구항 7

제 4항에 있어서,

상기 프로세서는 상기 출력 인터페이스를 통해 베이지안 신뢰구간 그래프를 제공하고, 상기 베이지안 신뢰구간 그래프는 상기 용량반응곡선에 대하여 베이지안 신뢰구간을 나타내는 것인, 용량반응곡선의 추정 및 기준용량을 계산하는 장치.

청구항 8

용량 반응곡선의 추정 및 기준용량을 계산하는 방법에 있어서,

분석대상에 대하여 사전 분포를 설정하는 단계;

용량반응곡선에 대하여 형태제약을 부과하는 단계;

상기 부과하는 단계를 통해 부과된 형태제약을 반영한 용량반응모형에 대하여 베이지안 방법론에 따라 사후분포를 계산하고

상기 계산하는 단계를 통해 계산된 사후분포로부터 임의표본을 추출하는 단계;

상기 추출하는 단계를 통해 추출된 상기 임의표본을 통해 용량반응곡선을 추정하는 단계; 및

상기 추정하는 단계를 통해 추정된 상기 용량반응곡선에 대하여 상기 기준용량을 계산하는 단계를 포함하며,

상기 분석대상은 비모수 함수 회귀모형으로 가정하는 것인, 용량반응곡선의 추정 및 기준용량을 계산하는 방법.

청구항 9

제 8항에 있어서,

상기 추정하는 단계는

사용자 인터페이스를 통해 사용자로부터 입력된 상기 분석대상에 대하여,

상기 사용자 인터페이스를 통해 사용자로부터 입력된 상기 형태제약을 고려하여 상기 용량반응곡선의 추정이 이루어지는 것인, 용량반응곡선의 추정 및 기준용량을 계산하는 방법.

청구항 10

제 8항에 있어서,

상기 기준용량을 계산하는 단계는

사용자 인터페이스를 통해 사용자로부터 입력된 상기 분석대상에 대하여;

상기 사용자 인터페이스를 통해 사용자로부터 입력된 기준용량계산 방법에 따라 용량의 범위를 등간격으로 나누어 격자를 생성하는 단계;

상기 생성하는 단계를 통해 생성된 격자에 따른 격자값에 대하여 상기 추정된 용량반응곡선상의 반응값을 계산하는 단계를 더 포함하고,

상기 기준용량은 상기 기준용량계산 방법에 따라 결정된 기준반응에 인접하는 상기 반응값에 해당하는 상기 격자값인 것인, 용량반응곡선의 추정 및 기준용량을 계산하는 방법.

청구항 11

제 8항에 있어서,

상기 사후분포를 계산하는 단계는,

상기 사후분포 중 완전 조건부 사후 분포에 대하여 깃스 표집기를 사용해 임의 표본을 추출하는 단계를 더 포함하고,

상기 추출하는 단계를 통해 추출된 임의 표본에 대하여 상기 기준용량을 계산하는 것인, 용량반응곡선의 추정

및 기준용량을 계산하는 방법.

청구항 12

제 8항에 있어서,

상기 사후분포를 계산하는 단계는,

상기 사후분포 중 완전 조건부 사후 분포를 정의하기 어려운 비모수 함수에 대하여 확률 보행 메트로폴리스-헤이스팅스 알고리즘을 사용해 임의 표본을 추출하는 단계를 더 포함하고,

상기 추출하는 단계를 통해 추출된 임의 표본에 대하여 상기 기준용량을 계산하는 것인, 용량반응곡선의 추정 및 기준용량을 계산하는 방법.

청구항 13

청구항 8 내지 청구항 12중 어느 한 항에 기재된 방법을 컴퓨터상에서 수행하기 위한 프로그램을 기록한 컴퓨터 판독 가능한 기록 매체.

발명의 설명

기술 분야

[0001] 본 발명은 용량반응곡선의 추정 및 기준용량 계산을 위한 장치 및 방법에 관한 것이다.

배경 기술

[0002] 용량반응자료를 사용한 기준용량의 계산에 있어서 우선적으로 고려해야 할 사항은 자료에 적합한 용량반응곡선을 추정하는 것으로, 기존에는 로지스틱(Logistic) 모형, 와이불(Weibull) 모형 및 힐(Hill) 모형과 같이 단조 증가 또는 감소하는 모수적(Parametric) 용량반응곡선을 사용해 왔다.

[0003] 하지만, 모수적 모형들은 비모수적 모형에 비해 적합도가 떨어지며, 해당 모형을 통해 기준용량을 계산할 경우, 실제 반응곡선이 해당 모수적 모형과 다를 확률이 높기 때문에 부정확한 모형 추정을 초래하게 되어 과소추정(underfitting) 또는 과다추정(overfitting)과 같이 정확한 기준용량 계산이 어렵게 된다.

[0004] 한편, 이러한 문제를 해결하기 위해 도입된 단조 증가 또는 감소하는 비모수 용량반응곡선의 추정방법이 연구되고 있으며, 이를 실제적으로 구현할 수 있는 소프트웨어가 등장했다.

[0005] 그러나 해당 소프트웨어는 단조 증가 혹은 감소하는 특정한 함수형태만을 사용하기 때문에, 용량과 반응의 관계가 추가적인 형태 제약(예를 들어, 오목/볼록 또는 S자 성장곡선의 형태)인 경우 추정에 있어 문제가 발생할 수 있다. 또한, 실제 응용연구에 있어 사용자들에게 익숙하지 않은 소프트웨어 프로그래밍 환경과 새롭게 배워야 하는 어려움 때문에, 실제 사용에 있어서 많은 제약이 발생할 수 있다.

[0006] 따라서, 형태제약 용량반응곡선의 추정과 기준용량 계산을 위한 새로운 알고리즘과 사용하기 쉬운 컴퓨터 소프트웨어의 개발이 필요한 실정이다.

[0007] 한편, 선행문헌인 Bornkamp, B. and Ickstadt, K. (2009). Bayesian Nonparametric Estimation of Continuous Monotone Functions with Applications to Dose-Response Analysis, Biometrics, 65.1, 198-205. 는 단조증가와 단조감소하는 형태제약을 반영한 비모수 함수의 용량 반응곡선의 추정방법에 대하여 개시하고 있다.

발명의 내용

해결하려는 과제

[0008] 본 발명은 분석대상을 비모수 함수 회귀모형으로 가정하고, 비모수 함수에 대하여 형태제약을 부과한 뒤 베이지안 방법을 적용하여 용량반응곡선을 추정하고 기준용량을 계산하는 장치 및 방법을 제공하는데 그 목적이 있다.

[0009] 다만, 본 실시예가 이루고자 하는 기술적 과제는 상기된 바와 같은 기술적 과제들로 한정되지 않으며, 또 다른 기술적 과제들이 더 존재할 수 있다.

과제의 해결 수단

[0010] 상술한 기술적 과제를 달성하기 위한 기술적 수단으로서, 본 발명의 일 실시예에 따른 용량반응곡선의 추정 및 기준용량을 계산하는 장치는 사용자의 입력값에 따라 상기 용량반응곡선을 추정하고, 상기 기준용량을 계산하는 프로그램을 저장하는 메모리; 및 상기 프로그램을 실행하는 프로세서를 포함하되, 상기 프로세서는 상기 프로그램의 수행에 따라, 분석대상에 대하여 사전분포를 설정하고, 용량반응곡선에 대하여 형태제약을 부과하고, 부과된 형태제약에 따른 용량 반응모형에 대하여 베이지안 방법론에 따라 사후분포를 계산하고, 상기 계산된 사후분포로부터 임의표본을 추출하고, 상기 추출된 임의표본을 통해 용량반응곡선의 추정하고, 추정된 상기 용량반응곡선에 대한 상기 기준용량을 계산한다. 이때, 분석대상은 비모수 함수 회귀모형으로 가정한다.

[0011] 또한, 본 발명의 일 실시예에 따른 용량 반응곡선의 추정 및 기준용량을 계산하는 방법은 분석대상에 대하여 사전 분포를 설정하는 단계; 상기 설정하는 단계를 통해 생성된 사전 분포에 대하여 형태제약을 부과하는 단계; 상기 부과하는 단계를 통해 부과된 형태제약을 반영한 용량반응모형에 대하여 베이지안 방법론에 따라 사후분포를 계산하는 단계; 상기 계산하는 단계를 통해 계산된 사후분포로부터 임의표본을 추출하는 단계; 상기 추출하는 단계를 통해 추출된 상기 임의표본을 통해 용량반응곡선을 추정하는 단계; 및 상기 추정하는 단계를 통해 추정된 상기 용량반응곡선에 대하여 상기 기준용량을 계산하는 단계를 포함한다. 이때, 분석대상은 비모수 함수 회귀모형으로 가정한다.

발명의 효과

[0012] 본 발명의 일 실시예에 따르면 용량반응곡선의 추정 및 기준용량을 계산하는 장치는 베이지안 방법을 사용함으로써, 기존의적합 방법보다 보다 정확하고 효율적일 수 있다. 더불어, 베이지안 방법을 사용하기 때문에, 소표본 자료에 있어도 효율적인 적합결과와 이에 따른 기준용량과 신뢰구간을 제공할 수 있다.

[0013] 또한, 본 발명의 일 실시예에 따르면 용량반응곡선의 추정 및 기준용량을 계산하는 장치는 사용자 인터페이스를 통해 입력되는 사용자의 입력값을 사용하여 계산이 이루어짐으로써, 통계방법론이나 컴퓨터 프로그래밍 분야의 전문가가 아니더라도 해당 분야의 실무자들 및 일반사용자도 편리하게 사용할 수 있다.

도면의 간단한 설명

[0014] 도 1은 본 발명의 일 실시예에 따른 용량반응곡선의 추정 및 기준용량을 계산하는 장치의 구성도이다.

도 2는 본 발명의 일 실시예에 따른 용량반응곡선의 추정 및 기준용량을 계산하는 방법에 대한 순서도이다.

도 3a내지 도 3c는 본 발명에서 제안되는 프로그램의 사용자 인터페이스의 일예를 나타내는 도면이다.

도 4a 내지 도 4c는 본 발명에서 제안되는 프로그램의 출력 인터페이스의 일예를 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0015] 아래에서는 첨부한 도면을 참조하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 본 발명의 실시예를 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0016] 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 소자를 사이에 두고 "전기적으로 연결"되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미하며, 하나 또는 그 이상의 다른 특징이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0017] 이하의 실시예는 본 발명의 이해를 돕기 위한 상세한 설명이며, 본 발명의 권리 범위를 제한하는 것이 아니다. 따라서 본 발명과 동일한 기능을 수행하는 동일 범위의 발명 역시 본 발명의 권리 범위에 속할 것이다.

[0018] 구체적 설명에 앞서, 본 발명에서는 분석대상에 있어 베이지안 푸리에 급수 비모수 회귀 모형을 적용하고 단조 증가/감소/오목/볼록/S자 형태로 조합될 수 있는 총 열가지의 회귀함수에 대한 형태적 제약을 반영하여 용량반응곡선을 추정하고 기준 용량을 계산한다.

- [0019] 구체적으로, 본 발명의 분석대상에 대하여 비모수 회귀 모형이 적용된다.
- [0020] 비모수 회귀 모형은 회귀 함수에 대한 특정한 가정이 필요 없고, 주어진 분석대상에 기초해 비교적 자유로운 모양을 가질 수 있다. 하지만, 비모수 회귀 모형은 회귀 함수에 대하여 특정한 형태를 가정하지 않기 때문에 다양한 형태의 회귀 곡선을 적합할 수 있는 장점을 갖고 있지만, 유한 차원의 모수 공간이 아닌, 연속함수 또는 미분 가능 함수와 같은 무한 차원의 함수공간에서 적합해야 한다는 단점이 있다.
- [0021] 즉, 본 발명의 추론대상인 모수는 확률변수이다. 추론 대상은 고정된 상수가 아니라, 불확실성을 갖는 확률변수이며, 모든 확률변수는 확률분포와 이에 따른 확률을 통해 설명할 수 있다는 베이지안 방법을 기반으로 통계적 추론을 실시할 수 있다.
- [0022] 다시 말해, 베이지안 방법론을 적용하여 용량반응곡선을 추론시 베이즈 방법론을 통해 얻어지는 사후분포를 바탕으로 비모수에 대한 점추정과 가설검정이 이루어질 수 있다.
- [0023] 즉, 본 발명에서 제안되는 푸리에 급수를 이용하는 베이지안 방법론은 비모수 함수에 대하여, 단조증가나 감소, 볼록/오목과 같은 형태제약을 쉽게 반영할 수 있어, 비모수 함수 적합에 우수함을 보인다.
- [0024] 또한, 베이지안 방법론은 대표본을 통해서 모수에 대한 신뢰구간을 점근적으로 얻어야하는 전통적 방법의 한계를 극복하며, 소표본의 경우에서도 사후분포를 통해 모수에 대한 베이지안 신뢰구간을 손쉽게 구할 수 있다. 즉, 모수가 특정한 구간 안에 들어갈 확률에 대해서 직관적으로 설명할 수 있다.
- [0025] 이하에서는 첨부된 도면을 통해 본 발명에서 제안하는 용량반응곡선 추정 및 기준용량 계산 기술에 대하여 자세히 설명하도록 한다.
- [0026] 도 1은 본 발명의 일 실시예에 따른 용량반응곡선 추정 및 기준용량 계산 장치의 구성도 이다.
- [0027] 본 발명의 일 실시예에 따른 용량반응곡선 추정 및 기준용량 계산장치(100)는 메모리(110), 프로세서(120)를 포함한다.
- [0028] 메모리(110)는 사용자의 입력값에 따라 용량반응곡선을 추정하고, 기준용량을 계산하는 프로그램(이하, ‘용량반응곡선 추정 및 기준용량 계산 프로그램’이라 지칭함)을 저장한다.
- [0029] 도시되진 않았지만, 용량반응곡선 추정 및 기준용량 계산 프로그램은 사용자에게 사용자 인터페이스를 제공한다. 이때, 사용자는 사용자 인터페이스를 통해 사용자의 입력값을 입력할 수 있다. 즉, 사용자 인터페이스는 단조증가, 단조감소, 오목, 볼록 및 S자의 형태제약을 선택할 수 있도록 하고, 사용자는 사용자 인터페이스에서 제공되는 형태 제약 중 어느 하나를 선택하여 입력할 수 있다.
- [0030] 사용자 인터페이스는 그래픽 인터페이스 형태로 제공된다. 이때, 사용자는 그래픽 인터페이스에서 제공하는 용량반응곡선을 추정하고 기준용량을 계산하기 위한 입력값 (예를 들어, 분석에 사용할 분석자료, 형태제약사항 (단조증가/단조감소/볼록/오목/S자형태곡선), 기준용량계산방법 등)을 선택하여 입력값을 입력할 수 있다. 이하에서 설명될 도 3a 내지 도 3c를 통해 본 발명에서 제안되는 사용자 인터페이스를 구체적으로 살펴본다.
- [0031] 용량반응곡선 추정 및 기준용량 계산 프로그램은 사용자 에게 출력 인터페이스를 제공한다. 이때, 사용자는 출력 인터페이스를 통해 프로그램 수행에 따른 결과를 제공받을 수 있다.
- [0032] 예를 들어, 출력 인터페이스는 사용자에게 용량반응곡선 및 기준용량 계산 결과에 대한 요약 통계량을 제공할 수 있다.
- [0033] 또한, 출력 인터페이스는 트레이스 플롯을 제공할 수 있다. 사용자는 트레이스 플롯을 통해 용량반응곡선 적합에 필요한 해당 모수의 사후분포로부터 생성된 임의 표본들의 수렴여부를 확인할 수 있다.
- [0034] 또한, 출력 인터페이스는 베이지안 신뢰구간 그래프를 제공할 수 있다. 사용자는 베이지안 신뢰구간 그래프를 통해 용량반응곡선에 대하여 베이지안 신뢰구간을 확인할 수 있다.
- [0035] 이하에서 설명될 도 4a 내지 도 4c를 통해 본 발명에서 제안되는 출력 인터페이스의 일예를 자세히 살펴본다.
- [0036] 프로세서(120)는 용량반응곡선 추정 및 기준용량 계산 프로그램을 실행한다.
- [0037] 구체적으로, 프로세서(120)는 프로그램 수행에 따라,
- [0038] 프로세서(120)는 프로그램의 수행에 따라, 분석대상에 대하여 사전 분포를 설정하고, 용량반응곡선에 대한 형태제약을 부과하고, 부과된 형태제약에 따른 용량반응모형에 대하여 베이지안 방법론에 따라 사후분포를

계산하고, 계산된 사후분포로부터 임의표본을 추출하고, 추출된 임의표본을 통해 용량반응곡선의 추정하고, 추정된 용량반응곡선에 대한 기준용량을 계산한다. 이때, 분석대상은 비모수 함수 회귀모형으로 가정된다.

[0039] 여기서, 형태제약을 부과하는 단계는 인터페이스의 형태제약 선택부에서 제공되는 회귀함수의 형태는 단조증가, 단조감소, 오목, 볼록 및 s자 형태로 조합될 수 있고 사용자는 인터페이스에서 제공되는 총 열가지 조합의 회귀 함수 중 한가지를 선택할 수 있다.

[0040] 도 2는 본 발명의 일 실시예에 따른 용량 반응곡선 추정 및 기준용량 계산방법에 대한 순서도이다.

[0041] 본 발명의 일 실시예에 따른 용량반응곡선 추정 및 기준용량 계산은 분석대상에 대하여 사전 분포를 설정하는 단계(S210), 형태제약을 부과하는 단계(S220), 사후분포를 계산하는 단계(S230), 사후분포로부터 임의표본을 추출하는 단계(S240), 용량반응곡선을 추정하는 단계(S250), 기준용량을 계산하는 단계(S260)를 포함한다.

[0042] 본 발명의 일 실시예에 따른 용량반응곡선 추정 및 기준용량 계산 장치는 메모리에 저장된 프로그램 실행에 따라 용량반응곡선을 추정하고 기준용량을 계산한다. 이때, 본 발명에서는 추정해야하는 모수가 고정된 상수가 아니라, 불확실성을 갖는 확률변수이다. 즉, 모든 확률변수는 확률분포와 이에 따른 확률을 통해 설명할 수 있다는 베이지안 관점을 기반으로 통계적 추론을 실시하며, 베이지 방법론을 통해 얻어지는 사후분포를 바탕으로 모수에 대한 점추정과 가설검정이 이루어진다.

[0043] 다시말해, 본 발명에서는 분석대상을 베이지안 푸리에 급수 비모수 회귀 모형으로 가정하고, 단조증가/감소/오목/볼록/S자 형태 등의 열가지의 회귀함수에 대한 형태적 제약을 반영한다. 이때, 회귀 함수의 도함수에 대한 모형적합을 고려하여 용량반응곡선을 추정하고 기준용량을 계산한다.

[0044] 설명에 앞서, 베이지 방법론은 상호배반이며 합집합이 전체집합인 사건들 ($B_1, \dots, B_k$)에 대해 다른 사건

A 가 일어났을 때 조건부 확률의 정의를 사용해 특정 사건 ($B_i, 1 \leq i \leq k$)가 일어날 확률을 계산

하며, 특정 사건 (C)가 일어날 확률을 $p(C)$ 라 표현했을 때, 하기의 [수학식 1]과 같이 표현될 수 있다.

수학식 1

$$p(B_j|A) = \frac{p(B_j)p(A|B_j)}{\sum_{i=1}^k p(A|B_i)p(B_i)}$$

[0045]

[0046] 즉, 본 발명에서 제안되는 푸리에 급수 비모수 회귀 모형으로 가정하여 베이지 방법론을 적용할 경우, ($x_1, \dots, x_n$)이라는 분석대상이 주어졌을시, ($x_i, 1 \leq i \leq n$)와 (θ)가 연속형 확률변수일

때 구하고자 하는 모수(θ)의 분포는 아래의 [수학식 2]와 같은 식을 통해 계산할 수 있다.

수학식 2

$$\pi(\theta | x_1, \dots, x_n) = \frac{f(x_1 | \theta) \cdots f(x_n | \theta) p(\theta)}{\int_{\Theta} f(x_1 | \theta) \cdots f(x_n | \theta) p(\theta) d\theta}$$

[0047]

[0048] 여기서, $p(\theta)$ 는 모수 (θ)에 대한 사전분포(prior distribution)를 나타내고, $f(x|\theta)$ 는 우도함수 (likelihood function)이며, $\pi(\theta | x_1, \dots, x_n)$ 는 사후분포(posterior distribution)를 나타낸다.

[0049] 예를 들어, 본 발명에서는 사전분포를 사용하는 회귀 모형의 사후분포를 계산하고, 메트로폴리스-헤이스팅스 알고리즘을 통해 임의의 난수를 추출 후 해당 난수들을 사용해 용량반응곡선의 추정 및 기준용량을 계산할 수 있다.

[0050] 이하에서는 도 2를 참고하고, 수학식을 통해 용량반응곡선 추정 및 기준용량 계산방법에 대해 구체적으로 설명하도록 한다.

[0051] 먼저, 프로그램 실행에 따라 사용자에게 의해 입력된 분석대상에 대하여 사전 분포를 설정한다(S210).

[0052] 본 발명에 있어서, 분석대상은 회귀 모형으로 가정된다. [수학식 3]은 분석대상에 대한 용량반응곡선의 추정을 위한 회귀 모형을 수학식으로 나타낸 것이다.

수학식 3

$$u = f(x) + \epsilon$$

[0053]

[0054] 여기서, 좌변의 (u)는 용량반응을 의미하며 $f(x)$ 는 형태제약이 부가된 비모수 함수를 의미한다.

[0055] 또한, $f(x)$ 는 특정한 함수 형태를 가정하지 않는 비모수 함수로서 $[0,1]$ 과 같이 주어진 유계 구간에서 (bounded interval)에서 미분 가능한 함수라고 가정할 수 있다.

[0056] 이때, 비모수적 미분 가능한 함수는 푸리에 기저의 무한 합으로 표현할 수 있다. [수학식 4]는 비모수적 미분 가능한 함수를 코사인(cosine) 기저함수(basis function)의 무한합으로 표현한 것을 나타낸다.

수학식 4

$$f(x) = \sum_{j=1}^{\infty} \theta_j \phi_j(x)$$

[0057]

[0058] 한편, 본 발명에서 제안되는 프로그램은 실제자료를 통한 모형적합에 있어서 유한한 (J)개의 코사인 함수를

사용하도록 할 수 있다. 여기서, (J)의 값은 사용자 인터페이스를 통해 사용자에게 의해 직접 입력될 수 있다.

[0059] 예를 들어, 일예에 따른 본 발명의 프로그램에서는 기본값으로 ($J=50$)이 설정되었으며, 사용자 인터페이스를 통해 사용자가 직접 해당 값을 조정할 수 있다.

[0060] 여기서, 사용자에게 의한 코사인 함수의 갯수인 (J)의 선택문제는 모형적합 전체에 있어서 중요한 값이다. 무한함으로 표현되는 [수학식 4]의 근사가 더 중요한 것으로 알려져 있으나, 일반적으로 (J)의 값이 커지게 되면, 모형 적합과정에서 다소 시간이 많이 걸릴 수 있다.

[0061] [수학식 5]는 구체적으로 사용하게 될 코사인 기저함수의 형태를 나타낸다. 여기서, $x_{(n)}$ 은 자료의 최댓값 $x_{(1)}$ 은 최솟값을 의미한다.

수학식 5

$$\phi_j(x) = \left(\frac{2}{x_{(n)} - x_{(1)}} \right)^{0.5} \cos \left(\pi j \left(\frac{x - x_{(1)}}{x_{(n)} - x_{(1)}} \right) \right)$$

[0062]

[0063] [수학식 4]에서 추정해야하는 모수들은 코사인 기저함수의 계수인 θ_i 이다.

[0064] 이때, 계층적 사전분포(hierarchical prior)를 사용할 수 있으며, 코사인 기저함수(θ_i)의 추정을 위하여 정규 사전분포를 사용할 수 있다. 예를 들어, τ 와 γ 라는 추가적인 초모수(hyperparameter)를 통해 $f(x)$ 의 적합도를 높일 수 있으며, 추가적인 초모수를 위한 사전분포를 사용하는 계층적 방식을 고려할 수 있다. 사

용된 사전분포들은 [수학식 6]과 같으며, 여기서 (σ^2)는 오차항의 분산이다.

수학식 6

$$\begin{aligned} \theta_j &\sim N(0, \tau^2 \exp(-\gamma j)), \tau^2 \sim IG\left(\frac{r_{0\tau}}{2}, \frac{s_{0\tau}}{2}\right), \\ \gamma &\sim \text{Exp}(w_0), \sigma^2 \sim IG\left(\frac{r_{0\sigma}}{2}, \frac{s_{0\sigma}}{2}\right) \end{aligned}$$

[0065]

[0066] 다음으로, 사전 분포를 설정하는 단계(S210)를 통해 설정된 사전 분포에 대하여 형태제약을 부과한다(S220).

[0067] 본 발명에서는 사용자 인터페이스를 통해 사용자가 형태제약을 선택할 수 있다.

[0068] 예를 들면, 사용자 인터페이스는 단조증가, 오목, 볼록, 변곡점을 기준으로 오목에서 볼록으로 변화하는 형태, 변곡점을 기준으로 볼록에서 오목으로 변화하는 형태와 단조감소 하며 오목, 볼록, 변곡점이 존재하는 두 가지 경우 등 총 10가지 형태제약을 제공하고, 사용자는 사용자 인터페이스에서 해당 형태제약 중 한가지를 선택할 수 있다.

[0069] 이때, 비모수 회귀 함수의 도함수의 성질을 사용할 수 있으며, 도함수가 양의 값일 때 단조 증가하고, 도함수가 음의 값일 때 단조 감소하는 성질을 사용하여 용량반응곡선을 추정할 수 있다. 즉, 비모수 회귀 함수

$$f(x)$$

에 대한 형태제약을 위해 비모수 회귀 함수의 도함수 $f'(x)$ 의 성질을 활용할 수 있다.

[0070] 다시 말해서, 종래에는 $f(x)$ 를 직접 푸리에 급수로 표현해 필요한 해당 모수들을 추정해왔지만, 본 발명에서

는 특정한 형태적 제약을 갖는 회귀 함수에 대한 추정을 위해 일계도함수 $f'(x)$ 혹은 이계도함수 $f''(x)$

보다 일반적으로 n 차 도함수를 먼저 고려한 후, 이를 적분하는 방식을 통해 원래의 함수 또는 역도함수(anti-derivative)를 추정하도록 하였다.

[0071] 예를 들어, 사용자가 사용자 인터페이스를 통해 단조증가, 단조감소인 형태제약을 선택할 경우, 단조 증가, 단

조감소의 용량반응곡선의 추정은, 비모수 회귀 함수의 도함수 $f'(x)$ 의 성질은 양(+)의 값 또는 음(-)의 값

만을 갖는다는 조건을 통해, 원래의 함수 $f(x)$ 가 단조증가인지 또는 단조감소인지를 결정될 수 있다.

수학식 7

$$\begin{aligned} f'(x) &= \frac{df(x)}{dx} = \left(\sum_{j=1}^J \theta_j \phi_j(x) \right)^2 \\ &= \sum_{j=1}^J \theta_j^2 \phi_j^2(x) + 2 \sum_{i=1}^{J-1} \sum_{k=i+1}^J \theta_i \theta_k \phi_i(x) \phi_k(x) \end{aligned}$$

[0072]

[0073]

[0074] 즉, 단조 증가, 단조감소의 용량반응곡선의 추정에 있어서, 특정 함수의 일계도함수가 음이 아닐 경우 ($f'(x) \geq 0$) 원래의 함수가 단조 증가한다는 성질을 사용하기 위해 추정코자 하는 비선형 $f(x)$ 의

일계도함수인 $f'(x)$ 를 푸리에 급수의 제곱꼴로 표현 가능하다 가정했으며, 이는 [수학식 7]과 같이 표현될 수 있다. 이는 단조 증가 혹은 단조 감소하는 용량반응곡선의 추정에 사용된다.

[0075] [수학식7]을 통해 $f'(x)$ 를 추정 후 $f(x) = f(0) + \int_0^x f'(t)dt$ 를 사용해 원래의 용량반응곡선을 추정할 수 있다.

[0076] 다른 예를들면, 사용자가 사용자 인터페이스를 통해 오목/볼록 형태제약을 선택할 경우, 오목(convex)한 용량반응곡선 $f(x)$ 의 이계도함수인 $f''(x)$ 를 푸리에 급수의 계수로 표현한 [수학식 8]을 사용해 모수를 추정할 수 있으며, 볼록(concave)한 함수의 경우 [수학식 8]의 우변에 -1 을 곱해 추정할 수 있다.

수학식 8

$$f''(x) = \frac{d^2 f(x)}{dx^2} = \left(\sum_{j=1}^J \theta_j \phi_j(x) \right)^2$$

[0077]

[0078] 또 다른 예를 들면, 사용자가 사용자 인터페이스를 통해 S자 형태의 형태제약을 선택할 경우, S자 형태인(변곡점을 갖고 있는) 용량반응모형 $f(x)$ 는 아래의 [수학식 9]와 같이 설정하였고, [수학식 9]의 (h)는 (

$$h(x) = \frac{1 - \exp[\psi(x - \omega)]}{1 + \exp[\psi(x - \omega)]}$$

$\psi > 0$ $0 < \omega < 1$)와 ()일 때 인 스퀴시

(squish)함수이며, (δ)와 (ζ)는 (1 또는 -1 로), 단조증가하며 변곡점 (ω)를 기점으로 오목한 형태에서 볼록한 형태로 바뀌거나, 단조감소하며 (ω)를 기점으로 볼록에서 오목으로 바뀌는 형태 등 총 네 가지의 서로 다른 형태의 용량반응모형을 표현할 수 있다

[0079] 예를 들어, [수학식 9]와 같은 용량반응모형 $f(x)$ 를 가정했을 때, 그것의 일계도함수 $f'(x)$ 와 이계도함수 $f''(x)$ 는 [수학식 10]과 같게 계산되며, ($x \leq \omega$)를 만족할 때 양이 아니며 ($x > \omega$)일 때 음이 아님을 알 수 있다. 즉, $f(x)$ 는 ($x \leq \omega$)에 대해 오목하며($x > \omega$)에 대해 볼록한 형태를 갖게 된다.

수학식 9

$$f(x) = \delta\zeta \int_0^x \int_0^s Z^2(t)h(t)dt ds - \delta\xi(x-0.5) + \alpha(x-0.5),$$

$$\xi = \min \left[0, \min_{x \in [0,1]} \zeta \int_0^x Z^2(s)h(s)ds \right].$$

[0080]

수학식 10

$$f'(x) = \delta\zeta \int_0^x Z^2(s)h(s)ds,$$

[0081]

[0082] 다음으로 형태제약을 부과하는 단계(S220)를 통해 부과된 형태제약을 반영한 용량반응모형에 대하여, 앞서 가정한 사전분포와 가능도를 바탕으로 베이지안 방법론에 따라 사후분포를 계산한다(S230). 다음으로, 사후분포로부터 임의표본을 추출하는 단계(S240)를 통해 추출된 임의표본을 통해 용량반응곡선을 추정한다(S250).

[0083] 이 경우, 상기 설정된 사전분포와 용량반응모형으로부터 베이지안 방법론을 적용해 유도된 사후분포는 정확한 계산이 어렵기 때문에 마르코프 연쇄 몬테카를로(Markov chain Monte Carlo, MCMC)방법을 통해서, 해당 사후분포로부터 임의표본(random sample, 난수)를 생성할 수 있다.

[0084] 본 발명의 일 실시예에 따른 용량반응곡선을 추정하고 기준용량을 계산하는 방법은 사후분포 중 완전 조건부 사후분포에 대하여 깃스 표집기를 사용해 임의표본을 추출하는 단계를 더 포함할 수 있다. 이때, 추출하는 단계를 통해 추출된 임의표본에 대하여 기준용량을 계산할 수 있다.

[0085] 구체적으로, 수학식을 통해 살펴보면, (σ^2, τ^2)의 경우, 임의표본을 생성하기 위한 완전 조건부 사후분포(full conditional posterior distribution)가 [수학식 11]과 같이 쉽게 계산되기 때문에 깃스 표집기(Gibbs sampler)를 사용해 사후표본을 추출할 수 있다.

수학식 11

$$\sigma^2 | \theta, \tau^2, \gamma, y \sim IG \left(\frac{n+r_{0\sigma}}{2}, \frac{s_{0\sigma}}{2} + \frac{1}{2} (y - f(x))^T (y - f(x)) \right)$$

$$\tau^2 | \theta, \gamma, y \sim IG \left(\frac{J+r_{0\tau}}{2}, \frac{s_{0\tau}}{2} + \frac{1}{2} \sum_{j=1}^J \frac{\theta_j}{\exp(-\gamma_j)} \right)$$

[0086]

[0087] 즉, 깃스 표집기는 조건부 확률 분포(conditional probability distribution)으로부터 순차적으로 표본을 추출해 원하는 확률변수들의 결합 확률 분포(joint probability distribution)을 생성하는 방법으로, 예컨대 k 개

$$X^{(0)} = (x_1^{(0)}, \dots, x_k^{(0)})$$

의 서로 다른 확률 분포로부터 표본을 생성하고자 할 때, 우선 초깃값을 선택한다. 그 뒤 한 개의 확률변수를 제외한 나머지 변수들로 이루어진 조건부 확률분포로부터 표본을 추출해

나가는 과정을 반복한다.

[0088]

다시말해, (t) 번째 추출의 경우 조건부 확률분포 $p(x_i^{(t)} | x_1^{(t)}, \dots, x_{i-1}^{(t)}, x_{i+1}^{(t-1)}, \dots, x_k^{(t-1)})$ 로

부터 $x_i^{(t)}$ 를 추출 후 $X^{(t)} = (x_1^{(t)}, \dots, x_k^{(t)})$ 를 표본으로 저장하여 $t+1$ 번째 난수를 생성하며, 위와 같은 난수 생성을 무수히 많이 반복한다.

[0089]

다른 예를 들어, 본 발명의 일 실시예에 따른 기준용량 계산 방법은 사후 분포 중 완전 조건부 사후 분포를 정의하기 어려운 비모수 함수에 대하여 확률 보행(random walk) 메트로폴리스-헤이스팅스 알고리즘을 사용해 임의의 표본을 추출하는 단계를 더 포함할 수 있다. 이때, 추출하는 단계를 통해 추출된 임의의 표본에 대하여 기준용량을 계산할 수 있다.

[0090]

구체적으로, (θ) 의 경우 완전조건부 사후분포를 정의하기 어렵기 때문에 확률보행(random walk) 메트로폴리스-헤이스팅스 알고리즘(Metropolis-Hastings algorithm, MH)을 사용해 표본을 추출할 수 있다.

[0091]

확률보행(random walk) 메트로폴리스-헤이스팅스 알고리즘의 경우 난수를 생성함에 있어 비례 인수(proportionality factor) (η) 에 의존하게 되고 여러 시행착오를 겪어 최적의 (η) 를 설정하게 되지만, 본 발명의 경우 연구자의 주관을 배제하며 보다 강건하게 난수 (θ) 를 생성하고자 (η) 를 역감마분포를 따르는 확률보행이라 가정한다. 그 후, (θ) 와 (η) 의 결합확률분포를 따르는 난수를 생성하며, 적분을 통해 (θ) 만을 남겨 분석에 활용할 수 있다. 이때, (t) 번째 난수 생성에는 [수학식 12]와 같은 제안분포를 사용한다.

수학식 12

$$\begin{aligned} \eta &\sim IG(a_T, b_{\eta, T}) \\ \theta_0 &\sim N(\theta_0^{(t)}, 5.66\eta\sigma v_{\theta}^2) I(\theta_0 \geq 0) \\ \theta_j &\sim N(\theta_j^{(t)}, 5.66\eta\sigma^2 \exp(-j\gamma)) \text{ for } j \geq 1 \end{aligned}$$

[0092]

메트로폴리스-헤이스팅스 알고리즘은 김스 표집기를 일반화한 알고리즘으로 완전 조건부 사후분포를 특정 분포로 정의하기 어려울 때 사용되며, 일정 조건을 만족하는 제안분포로부터 생성된 난수가 [수학식 13]과 같은 조건을 만족했을 경우 해당 난수를 채택, 그렇지 않으면 기각하며, 채택된 난수들은 원하는 목표분포로부터의 난수가 되기 때문에, 해당 표본들을 분석에 활용할 수 있다.

[0094]

즉, (t) 번째 반복의 경우의 제안분포를 $q(\cdot | X^{(t)})$ 라 했을 때, 제안분포로부터의 난수 Y 를 생성한 후,

$X^{(t+1)} = Y$ $X^{(t+1)} = X^{(t)}$ [수학식 13]과 같은 조건을 만족했을 경우, 를 채택, 그렇지 않으면 로 두는 반복을 목표분포에 수렴할 때까지 반복하게 된다. [수학식 13]의 U는 0과 1사이의 균일분포를 따르며 매 반복마다 $X^{(t)}$ 와 함께 생성되고, $f()$ 는 원하는 목표분포를 나타낸다.

수학식 13

$$U \leq \frac{f(Y)g(X^{(t)}|Y)}{f(X^{(t)})g(Y|X^{(t)})}$$

[0095]

[0096] 즉, 메트로폴리스-헤이팅스 알고리즘 및 깃스 표집기를 사용해 획득한 임의 표본에 대해 기준 용량을 계산할 수 있다. 이하에서는 기준용량을 계산하는 단계에 대하여 자세히 설명하도록 한다.

[0097] 다음으로, 용량반응곡선을 추정하는 단계(S250)를 통해 추정된 용량반응곡선의 기준용량을 계산한다(S260). 예를 들어, 사용자 인터페이스를 통해 사용자로부터 입력된 상기 분석대상에 대하여; 사용자 인터페이스를 통해 사용자로부터 입력된 기준용량계산 방법에 따라 용량의 범위를 등간격으로 나누어 격자를 생성하는 단계, 생성하는 단계를 통해 생성된 격자에 따른 격자값에 대하여 상기 추정된 용량반응곡선상의 반응값을 계산하는 단계를 더 포함할 수 있다.

[0098] 이때, 기준용량은 상기 기준용량계산 방법에 따라 결정된 기준반응에 인접하는 상기 반응값에 해당하는 상기 격자값일 수 있다.

[0099] 구체적으로, 기준용량(Benchmark dose, BMD)은 기준반응(Benchmark Response)이라고 불리는 특정 위험수준에 대응하는 용량으로, 용량반응곡선의 추정 후 간단한 계산을 통해 구할 수 있으며, 반응변수가 이산형 확률변수일 경우와 연속형 확률변수일 경우 각기 다른 방식을 통해 계산된다. 이산형 반응변수의 경우 용량반응곡선은 용량에 따른 반응확률로 추정되기 때문에 기준반응이나 부작용 반응이 발생할 확률에 대응하는 용량이 기준용량이 되지만, 본 발명에서는 연속형 반응변수에 대한 기준용량만을 고려한다.

[0100] 예를들어, 연속형 반응변수에 대한 기준용량을 계산하기 위한 방법으로 는Crump(2002)의 방법, Ritz(2013)의 방법, BMDS(2014)의 방법이 있다.

[0101] 먼저, Crump(2002)의 방법은 사용자가 입력하는 절단값을 사용해 계산할 수 있다. [수학식 14]는 기준반응을

나타낸다. 이때, 푸리에 급수를 사용한 비모수 회귀분석을 통해 추정한 함수를 $f(x)$, 용량(x)에 대응하는 반응변수 (Y_x)가 평균이 $f(x)$ 이고 분산이 σ^2 인 정규분포를 따른다는 가정하에 (Y_x)의 사용자가 원하는 절단값 c 이하일 확률은 [수학식 15]와 같이 계산할 수 있다. 여기서, (Φ)는 정규분포의 누적확률밀도함수를 뜻한다.

수학식 14

$$BMR = \frac{(1 - P(BMD)) - (1 - P(\text{기본용량}))}{P(\text{기본용량})}$$

[0102]

수학식 15

$$P(x) \equiv P(Y_x < c) = \Phi\left(\frac{c - f(x)}{\sigma}\right)$$

[0103]

[0104] 여기서, 로그 변환된 설명용량($\log(x)$) 및 용량반응을 분석에 사용하므로 위의 기준용량을 사용할 경우 $P(\log(BMD))$ 는 [수학식 16]과 같이 표현되기 때문에 두 수식을 방법론할 경우 [수학식 17]로 나타낼 수 있다. 이 형태를 갖는 기준용량은 Crump의 하이브리드 계산법(hybrid-approach)이라 불린다.

수학식 16

$$P(\log(BMD)) = (1 - BMR)\Phi\left(\frac{\log(c) - f(\log(\text{기본용량}))}{\sigma}\right) \equiv p^*$$

[0105]

수학식 17

$$BMD = \exp\{f^{-1}(\log(c) - \sigma\Phi^{-1}(p^*))\}$$

[0106]

[0107] 여기서, $f^{-1}(\log(c) - \sigma\Phi^{-1}(p^*))$ 를 계산함에 있어 비모수적 모형은 그 역함수를 해석적으로 구할 수 없기 때문에 (x)의 범위를 200등분해 총 200개의 x 를 우선적으로 계산하였고, 사용자가 입력하는 절단값 (c), 확률 (p^*)와 추정되는 (σ)인 ($\hat{\sigma}$)를 사용해 $\log(c) - \hat{\sigma}\Phi^{-1}(p^*)$ 를 계산한 후 해당 값과 가장 가까운 $f(x)$ 를 갖는 (x)값으로 외삽하는 방법을 사용할 수 있다.

[0108] 다른 예들들면, Ritz(2013)의 방법은 사용자가 입력한 기준위험을 사용해 계산된다.

[0109] 해당 방법은 기준용량을 계산함에 있어 크럼프(Crump)의 하이브리드 계산법과 유사하지만, 절단값을 고려하지 않는다. 즉, [수학식 17]에서는 $f^{-1}(\log(c) - \sigma\Phi^{-1}(p^*))$ 를 계산하지만 본 방법에서는 [수학식 18]을 사용해 기준용량을 계산하게 된다. 이 역시 해석적인 해를 구할 수 없기 때문에 200등분된 범위 중 가장 가까운 곳에 위치한 용량을 기준용량으로 구하는 방법이 사용되었다.

수학식 18

$$BMD = f^{-1}(\sigma\Phi^{-1}(p^*))$$

[0110]

[0111] 마지막으로 BMDs(2014)의 방법은 사용된 방법은 사용자가 입력한 기준반응인자(benchmark response factor)를 사용해 기준용량을 계산하는 방법이다. 이때 위험 수준의 종류를 설정하는 여러 방법이 존재하지만 본 발명에서는 표준편차 기반 위험 수준의 계산법이 사용됐고, 이는 [수학식 19]와 같이 표현될 수 있다. 여기서 BMRF는 기준반응인자를 뜻한다.

수학식 19

[0112]

$$BMR = f(0) + BMRF \times \sigma$$

[0113] 이때, 기준용량은 [수학식 20]와 같이 표현될 수 있으며, 첫 번째 방법과 마찬가지로 200등분된 범위에서 가장 가까운 값을 찾는 방법을 사용했다.

수학식 20

[0114]

$$\begin{aligned} f(BMD) - f(0) &= BMRF \times \hat{\sigma} \\ f(BMD) &= f(0) + BMRF \times \hat{\sigma} \\ BMD &= f^{-1}(f(0) + BMRF \times \hat{\sigma}) \end{aligned}$$

[0115] 도 3a내지 도 3c는 본 발명에서 제안되는 사용자 인터페이스(301)의 일예를 나타내는 도면이다. 도 3a 내지 도 3c는 용량반응곡선의 추정 및 기준용량을 계산하기 위하여 사용자에게 의한 입력값을 입력하기 위해 제공되는 그래픽 인터페이스를 나타낸다.

[0116] 도 3a를 살펴보면, 사용자는 사용자 인터페이스의 브라우저(Browse)(301)를 클릭하여 분석대상 파일을 입력할 수 있다. 또한, 초기값(Initialization)(305a), 소각과정을 위한 값(burn-in)(305b), 씨닝과정을 위한 값(Thin)(305c), 분석값(305d)을 입력할 수 있다.

[0117] 예를들어, 본 발명에서 사용되는 깃스 표집기와 메트로폴리스-헤이스팅스 알고리즘은 초기값(305a)에 영향을 받을 수 있기 때문에 일정 수준의 반복으로 얻는 표본들은 분석에 사용되면 안되므로 소각과정(burn-in)을 위한 값(305b)을 입력해야 한다.

[0118] 만약 분석에 활용할 난수가 총 10,000개일 경우 소각과정(burn-in)을 5,000으로 설정한다면 총 15,000개의 난수를 생성해 앞의 5,000개의 난수를 배제한 후 나중에 생성되는 10,000개의 난수만을 분석에 사용하게 된다. 또한 두 방법은 직전에 추출된 표본에 영향을 받아 자기상관(Autocorrelation)을 보일수 있기 때문에 연속적으로 생성된 표본 역시 올바른 분석을 방해할 수 있으므로 생성된 모든 난수를 분석에 사용하지 않고 매 n번째의 난수만을 분석에 활용하는 씨닝과정(Thin)을 거쳐야 한다. 즉, 총 10,000개의 난수를 분석에 사용하고자할 때, 씬(Thin)을 10으로 설정한다면 총 100,000개의 난수를 생성한 후 10, 20, ..., 99,990, 100,000번째의 난수만을 분석에 활용하게 된다.

[0119] 따라서, 본 발명은 초기값, 소각과정을 위한 값(burn-in), 씨닝과정을 위한 값(Thin), 분석값 (=분석에 사용할 표본의 개수)를 사용자로 하여금 직접 입력할 수 있다.

[0120] 도 3b를 살펴보면, 용량반응곡선의 가능한 형태제약의 종류를 선택하는 인터페이스를 나타낸 도면이다. 사용자는 형태제약(shape)(303)을 클릭하여 형태 제약의 종류를 선택할 수 있다.

[0121] 예를들어, 형태제약의 종류로는 자유형 곡선(free), 단조형 곡선(monotone), 볼록 곡선(convex), 오목 곡선(concave), 볼록에서 오목곡선(convex to concave), 오목에서 볼록곡선(concave to convex)등을 선택하여 입력할 수 있다

[0122] 더불어, 용량반응곡선은 일반적으로 주어진 용량에 대하여 반응곡선이 단조로 알려져 있고, 단조 함수의 경우는 기본적으로 증가하거나 감소하는 두가지 형태로 구분할 수 있다. 구체적으로 증가(감소)하는 반응곡선은 다시 각각 볼록한 형태로 증가하거나 오목한 형태로 증가한다고 가정할 수 있다.

[0123] 따라서 증가 또는 감소하는 두 가지 함수형태에 추가적으로 볼록/오목의 두 가지 제약조건을 추가하면 6가지의

형태제약이 고려될 뿐 아니라 증가하는 함수는 특정한 값에서 볼록함수에서 오목함수로 변하게 되는 변곡점을 갖게 되는 S자 형태를 갖게 되며, 이와 같이 변곡점을 포함하는 증가 또는 감소함수인 S자 곡선을 고려하면 4가지 추가적인 형태제약을 고려할 수 있다.

- [0124] 특히 S자 형태의 곡선은 용량반응곡선에서 대표적으로 사용할 수 있는 형태적 제약이라 할 수 있다.
- [0125] 도 3c는 기준용량을 계산하는 방식을 선택하는 화면을 나타낸 도면이다.
- [0126] 구체적으로, 사용자는 기준용량계산(BMD computation)탭을 클릭하여 절단값(cutoff), 기준위험(BMR), 기준반응인자(BMRF)등을 선택하여 사후표본에 대한 기준용량을 계산할 수 있다. 도 4a 내지 도 4c는 본 발명에서 제안되는 출력 인터페이스의 일예를 나타내는 도면이다.
- [0127] 도 4a 를 살펴보면, 출력 인터페이스는 용량반응곡선의 추정 및 계산된 기준용량에 대한 요약 통계량을 제공한다. 도 4a를 살펴보면, 도면에 모형 선택의 기준이 될 수 있는 가능도 통계량(Likelihood Statistics), 추정된 모수($\beta, \sigma, \alpha, \psi, \omega, \tau, \gamma, \zeta$)의 평균 및 표준편차 등의 통계량을 확인할 수 있다.
- [0128] 도 4b를 살펴보면, 출력 인터페이스는 추정된 모수들의 수렴진단을 할 수 있게 해주는 트레이스 플롯(trace plot)을 제공한다.
- [0129] 예를 들어, 사용자는 트레이스 플롯을 확인함으로써, 깁스 표집기(혹은 MH 알고리즘)를 통해 사후표본으로부터 추출된 난수가, 해당 사후분포로부터 추출된 것이 맞는지, 혹은 자기상관이 존재하는지 여부 등을 탐색적으로 판단할 수 있다.
- [0130] 도 4b를 살펴보면, 총 여섯종류의 트레이스 플롯을 확인할 수 있다.
- [0131] 도 4b의 (a)내지 (f)의 플롯들은 플롯들의 제목에 해당하는 모수들(예를들어, Beta, Sigma, Zeta, Tau, Gamma, Theta등)의 사후분포로부터 추출된 난수에 대한 수렴여부를 시각적으로 진단할 수 있다..
- [0132] 도 4c를 살펴보면, 출력 인터페이스는 사용된 자료와 추정된 용량반응곡선 및 곡선의 95% 베이지안 신뢰구간을 표현해주는 화면으로, 사용자가 직관적으로 해당 결과를 확인할 수 있다.
- [0133] 도 4c를 살펴보면, 용량반응곡선(412)은 사후분포로부터 추출된 임의표본을 사용해 계산된 회귀함수들의 평균반응곡선이다. 도 4의 95% 베이지안 신뢰구간(410a, 410b)은 용량반응곡선(412)의 추정에 사용된 1,000개의 회귀함수 중 상/하위 2.5% 분위수(quantile)의 값을 갖는 회귀함수를 사용하여 나타낸 베이지안 신뢰구간이고, 추정된 모수(400)는 관측값을 나타낸다. 즉, 본 발명은 프로그램 실행에 따라 획득된 용량반응곡선(412)에 대하여 신뢰구간(410a, 410b)에 따른 관측값(400)의 위치를 한눈에 확인할 수 있다.
- [0134] 예를 들어, 총 1,000개의 사후표본이 추출되었다면 해당 표본들을 사용해 각기 대응되는 1,000개의 회귀함수를 추정할 수 있으며, 1,000개의 회귀함수의 평균으로부터 계산된 용량반응곡선을 상정할 수 있다. 한편, 사용자의 신뢰구간 설정에 따라, 통계학 및 통계를 활용한 각종 응용 분야에서 일반적으로 유의수준을 5%로 분석 및 검정을 진행할 수 있으며, 대안적으로 90% 또는 99%등의 값을 설정하여 그 결과를 확인할 수도 있다.
- [0135] 이상에서 살펴본 바와 같이, 본 발명은 베이지안 형태제약 비모수 회귀모형을 이용한 기준용량 계산 시스템의 발명은 용량반응곡선 및 기준용량의 추정에 있어서 용량반응곡선의 특정한 형태 제약에 상관없이, 자동적으로 용량반응곡선을 추정해주며 아울러 베이지안 방법론을 이용한 적합방법을 통해, 기존의 용량반응적합 방법보다 우수한 성능과 보다 정확하고 효율적인 기준용량 계산방법을 제공한다.
- [0136] 또한, 본 발명은 특정한 용량반응모형에 대한 설계 및 계산을 위한 비용과 시간을 절감할 수 있으며, 해당 모형의 적합도와 관련된 각종 통계적 수치들을 제공함으로써 해당 추정의 유의성에 대한 추가적인 계산의 노력 또한 절감할 수 있게 한다.
- [0137] 또한 그래픽 사용자 인터페이스를 통해 계산이 이뤄지기 때문에, 통계방법론이나 컴퓨터 프로그래밍 분야의 전문가가 아니더라도 쉽게 사용이 가능한 소프트웨어를 제공함으로써, 해당 분야의 실무자나 일반적인 사용자도 편리하게 사용할 수 있다.
- [0138] 더불어, 베이지안 방법을 사용하면, 기존의 방법보다 모수의 추정에 있어서의 불확실성을 사후확률을 통해 보다 효율적으로 반영할 수 있다. 특히, 기존의 방법에서 표본의 크기가 큰 경우 기준용량에 대한 신뢰구간을 점근적 정규성을 바탕으로 계산하는 방식에 비해, 몬테카를로 방법(Monte Carlo method)을 통한 베이지안 추론을 통해 보다 효율적으로 신뢰구간을 계산할 수 있을 뿐 아니라 표본의 크기와 무관하게 계산할 수 있게 된다.

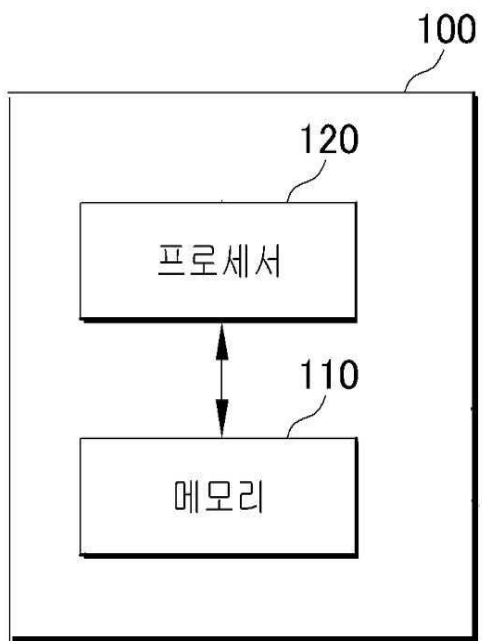
- [0139] 예를 들어 기준용량의 하위 95% 값을 의미하는 BMDL5 등의 계산에 있어, 베이지안 신뢰구간을 계산하는 방법을 사용할 수 있기 때문에 보다 직관적인 계산을 할 수 있다.
- [0140] 또한, 본 발명은 비모수적 회귀모형을 사용해 용량반응곡선을 추정하기 때문에 특정한 용량반응곡선의 가정이 필요하지 않고, 무엇보다도 용량반응곡선의 형태적 제약을 반영할 수 있다. 게다가 본 발명은 베이지안 방법을 사용하기 때문에, 소표본 자료에 있어서도 효율적인 적합결과와 이에 따른 기준용량과 신뢰구간을 제공할 수 있다.
- [0141] 설명된 용량 반응 곡선 추정 및 기준용량 계산 장치 및 방법은 컴퓨터에 의해 실행되는 프로그램 모듈과 같은 컴퓨터에 의해 실행 가능한 명령어를 포함하는 기록 매체의 형태로도 구현될 수 있다. 컴퓨터 판독 가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 휘발성 및 비휘발성 매체, 분리형 및 비분리형 매체를 모두 포함한다. 또한, 컴퓨터 판독가능 매체는 컴퓨터 저장 매체 및 통신 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함한다. 통신 매체는 전형적으로 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈, 또는 반송파와 같은 변조된 데이터 신호의 기타 데이터, 또는 기타 전송 메커니즘을 포함하며, 임의의 정보 전달 매체를 포함한다.
- [0142] 진술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.
- [0143] 본 발명의 범위는 상세한 설명보다는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

부호의 설명

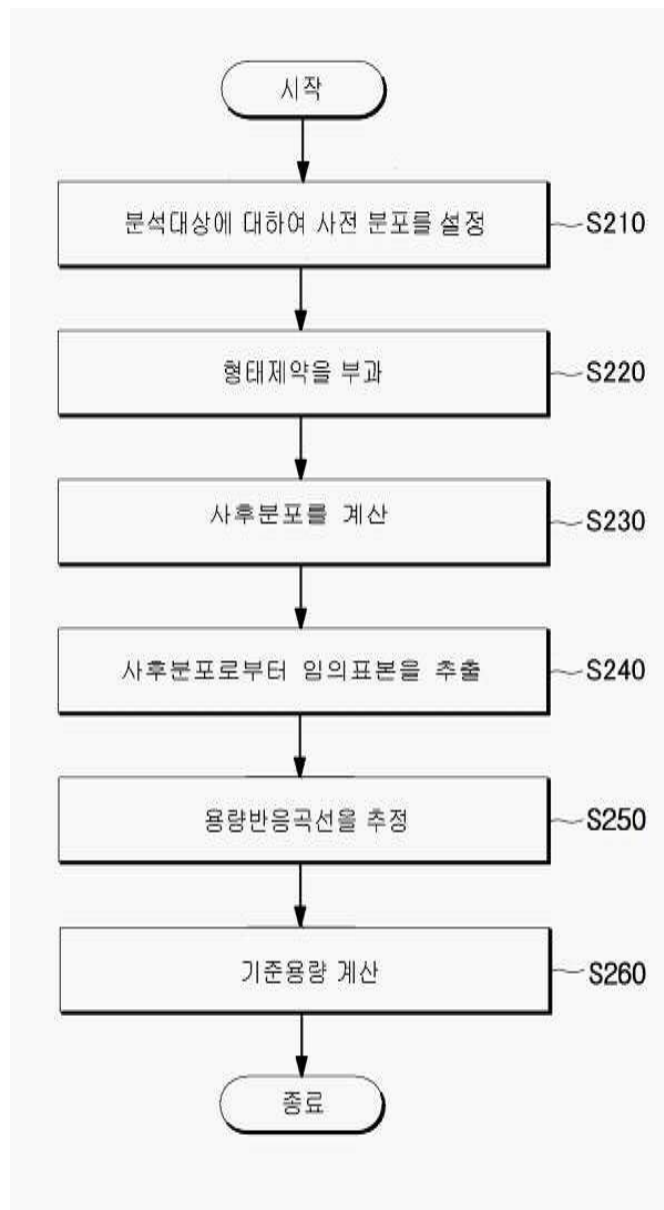
- [0144] 100: 용량반응곡선의 추정 및 기준용량을 계산하는 장치
110: 메모리
120: 프로세서

도면

도면1



도면2



도면3a

BSAR Graphical User Interface 1.1

File Help

Simulation Control Panel

File Name Browse...

☐ Increasing ☐ Decreasing

Shape

Number of basis

Initialization

Burn-In

Thin

For Analysis

BMD Computation

Method

Run Cancel

도면3b

BSAR Graphical User Interface 1.1

File Help

Simulation Control Panel

File Name

☒ Increasing ☐ Decreasing

Shape 303

Number of basis

Free

Monotone

Convex

Concave

Convex to Concave

Concave to Convex

Initialization

Burn-In

Thin

For Analysis

BMD Computation

Method

도면3c

BSAR Graphical User Interface 1.1

File Help

Simulation Control Panel

File Name

☒ Increasing ☐ Decreasing

Shape

Number of basis

Initialization

Burn-in

Thin

For Analysis

BMD Computation 304

Method

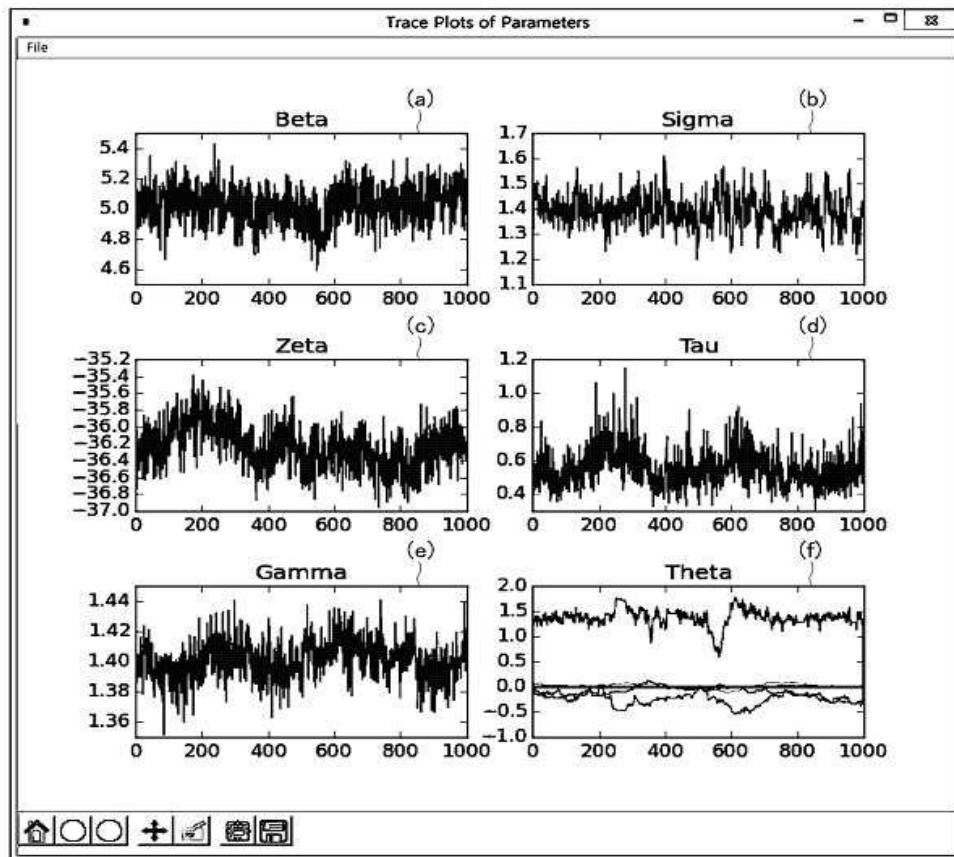
BMR

BMRF

도면4a



도면4b



도면4c

