

혼합순환분포를 이용한 유적 및 위상차 자료의 모형화 분석

장영미, 나종화

corolla@cdc.go.kr

질병관리본부 만성병조사팀, 충북대학교 정보통계학과

서울특별시 은평구 통일로 194

충청북도 청주시 흥덕구 성봉로 410(개신동) 충북대학교

02-380-2935

키워드 : 모형화, 순환자료, 다봉형, 최우추정법, EM

- Abstract -

일반적인 연속형 자료는 크기와 방향의 정보를 포함하는 것이 보편적이다. 그러나 실제 문제에서 방향의 정보에만 초점을 두어 각 또는 방위로 관측하는 경우가 종종 발생한다. 이러한 자료를 방향자료(directional data)라고 하며 특히, 원상에 표현되는 자료를 순환자료(circular data)라고 한다. 순환자료는 생물학을 비롯한 지질학, 경제학, 의학, 기상학, 물리학 등 다양한 분야에서 산출되고 있다. 위와 같이 각 또는 방위로 측정하여 생성될 수도 있으며 월 또는 시간과 같은 주기를 각 정보로 변환하여 생성할 수도 있다.

순환자료의 모형화 분석과 관련하여, 지금까지의 연구는 von Mises 분포를 중심으로 하는 단봉형의 대칭자료를 중심으로 진행되어 왔다. 본 논문에서는 실제의 자료분석 과정에서 매우 빈번하게 발생하는 이봉형 또는 다봉형의 순환자료에 대한 적합을 다루기로 한다. 순환분포의 혼합모형에 대한 모수추정은 분포함수의 복잡성으로 인한 계산상의 제약으로 인해 혼합-von Mises(von Mises Mixture) 분포를 중심으로 제한적으로 이루어져 왔다. 혼합-von Mises 분포의 최우추정에 대한 연구로 Jones and James(1972)가 이봉형의 방향자료에 대해 최대경사법을 사용하여 모수를 추정하였으며, Mardia and Sutton(1975)이 역시 이봉형(bimodal)에 국한된 혼합-von Mises 분포에 대해 Newton 알고리즘을 사용한 모수추정을 제안하였다.

본 논문에서는 일반적인 다봉형의 순환자료에 대한 적합모형으로 혼합-von Mises 분포를 제안하고, 이에 대한 최우추정법으로 EM알고리즘을 제시하고자 한다. 또한 이를 이용하여 다봉형의 가마터 유적자료와 위상차 자료의 모형화 분석을 수행하기로 한다.