

Analysis of Data from USA Region Using Additive Regression with Hilbertian Responses

기도형¹, 유명훈¹, 이경원², and 이동현²

¹*Department of Mathematics, Seoul National University*

²*Department of Statistics, Seoul National University*

December 13, 2019

1 Introduction

2 Data

분석에 이용한 자료는 복수의 채널을 통하여 획득하였다. 먼저, 인종의 경우 2017년에 각 주 별 인종 분포를 (1) 백인과 그 혼혈, (2) 흑인과 그 혼혈, (3) 아시안과 그 혼혈, (4) 아메리카 원주민과 알래스카 원주민, 그리고 그들의 혼혈, (5) 하와이 원주민과 그 혼혈로 나누어 조사한 자료를 이용하였다. 이 자료에서는 한 사람이 여러 인종의 혼혈인 경우, 그 사람을 해당되는 모든 항목에서 셈하고 있다. 우리는 분석과 시각화의 편의를 위하여 가장 많은 비율을 차지하는 (1) 백인과 그 혼혈, (2) 흑인과 그 혼혈, (3) 아시안과 그 혼혈을 분석에 이용하였고, 세 인종의 합을 전체로 하여 비율을 다시 계산하고 사용하였다.

다음으로 소득의 경우 2017년 각 주 별 일년 간의 가계 소득 분포의 중앙값 자료를 사용하였고, 연령의 경우 2017년 각 주 별 연령 분포의 중앙값을 사용하였다. 또한, 범죄율의 경우 2017년 각 주 별 인구 10만명 당 강력 범죄의 빈도 자료를 사용하였다. 마지막으로, 기후의 경우, 다른 자료들과 달리 연간 편차가 심하다는 점을 고려하여, 1981년부터 2010년까지의 각 주 별 평균 기온 자료와 평균 강수량 자료를 사용하였다.

3 Methods

4 Results

5 Conclusion

References