

10장. 보험료를 차등화하는 이유와 방법

◆ Introduction

자동차 보험의 경우, 연령, 차종, 주행거리, 사고이력 등에 따라 보험료가 다르게 책정된다. 보험료를 차등화 하는 이유는 위의 정보가 사고 건수나 손해액에 큰 영향을 미칠 수 있기 때문이다. 보험회사는 피보험자의 위험 정보를 파악하여 예상치 못한 손실에 대비해야 한다. 이번 장에서는 보험회사가 피보험자로부터 주어지는 정보를 반영하는 방법 중 하나인 베이지안 기법에 대해서 알아보겠다.

전통적인 점 추정, 구간추정, 가설검정에서는 Frequentist Approach 을 사용한다. 이 접근은 고정된 모수를 점이나 구간 같은 변하지 않는 것들로 추정한다. 이와 달리 Bayesian Approach은 모수를 확률분포의 형태로 추정하며, 이 확률분포는 경험 데이터를 통해 바뀐다고 생각한다. 새로 들어온 경험 데이터를 통해 모수에 대한 믿음(분포)를 계속 변형시켜 나간다고 생각하자.

베이지안의 구조를 이해하기 위해서 배워야 할 개념에는 사전 분포(prior distribution), 모델 분포(model distribution), 사후 분포(posterior distribution), 예측 분포(predictive distribution) 가 있고 이들의 의미와 관계를 명확하게 이해해야 한다. 사전 분포는 새로운 데이터를 관측하기 전에 특정 모수 θ 에 관한 주관적, 사전적 의견 혹은 믿음을 나타낸다. 이 글에서는 모수 θ 에 대한 사전분포를 $\pi(\theta)$ 로 표현하겠다. 다음으로 모델 분포는 모수 θ 가 주어졌을 때, 수집된 데이터 혹은 관측값들의 확률 분포이다. 모델 분포의 확률밀도함수를 $f_{X|\theta}(x|\theta)$ 라고 하자. 이 개념들을 통해 joint distribution과 marginal distribution 을 정의할 수 있다.

Joint Distribution $f_{X,\theta}(x, \theta) = f_{X|\theta}(x|\theta)\pi(\theta).$

Marginal Distribution $f_X(x) = \int f_{X|\theta}(x|\theta)\pi(\theta) d\theta.$

사후분포는 새로 관측된 데이터를 반영한 모수 θ 의 조건부 확률 분포이다. 수집된 데이터에 의해서 우리가 사전에 정의한 θ 에 대한 분포가 바뀐다. 조건부 확률분포로 표현한 사후 분포는 다음과 같다.

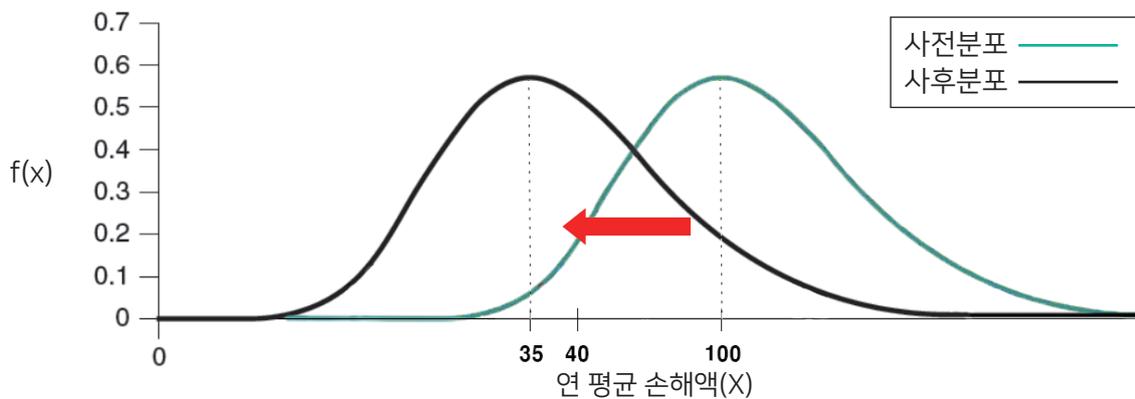
Posterior Distribution $\pi_{\theta|X}(\theta|X) = \frac{f_{X|\theta}(X|\theta)\pi(\theta)}{\int f_{X|\theta}(X|\theta)\pi(\theta) d\theta}.$

10장. 보험료를 차등화하는 이유와 방법

예측 분포는 데이터 \mathbf{X} 가 주어졌을 때에 미래에 관측될 y 의 조건부 확률 분포를 말한다. n 개의 관측치 $\mathbf{X} = [X_1 X_2 X_3 \cdots X_n]'$ 가 있을 때 $y(X_{n+1})$ 의 분포를 예측하는 것인데, 이때 θ 는 사후분포를 따른다.

$$\text{Predictive Distribution} \quad f_{Y|X}(y|\mathbf{X}) = \int f_{Y|\theta}(y|\theta) \pi_{\theta|\mathbf{X}}(\theta|\mathbf{X})d\theta$$

배운 개념을 가지고 간단한 예시를 들어보겠다. 보험회사가 보험 심사(Underwriting) 과정에서 피보험자 A를 고위험군으로 판단한 후 A의 모수에 대한 사전분포를 설정했다. 그리고 그 분포의 평균값(연 평균 손해액)이 100만원이라고 판단을 했다. 그리고 실제로는 알 수 없지만 A의 실제 연 평균 손해액은 40만원이라고 하자. (즉, A의 모수는 40만원이다.) 그러나 A는 최근 3년 동안 적은 손해액을 발생시켰다. 보험회사는 A의 경험 데이터를 반영하여 A의 모수에 대한 분포를 변경하고 평균 손해액을 35만원으로 수정한다. 이러한 절차가 반복되면 결국 A의 참 모수인 40만원으로 점진적으로 다가가게 된다.



베이저안 분석 기법은 모수 공간에 대해서 여러 차례 적분이나 합(summation)을 요구한다. 이 계산은 상당히 복잡하고 심지어 컴퓨터로도 해결이 불가능하다. 이러한 문제를 해결하기 위해 많은 연구가 진행되었고 그 결과 오늘날에 주로 널리 사용되는 방법은 MCMC(Markov Chain Monte Carlo) 시뮬레이션이다. 이에 관해 더 자세하고 깊은 내용이 알고 싶다면 관련 책을 참고하기 바란다.

★ 쉬어가는 페이지 ★



사진출처 : 뉴제주일보

“자동차 보험금 노린 ‘나이롱 환자’ 꼼짝 마!”

보험개발원은 자동차보험의 평균보험금 증가 주범으로 꼽히는 경상환자의 과잉치료비 지급 관련 제도 개선을 본격화한다고 2일 밝혔다. 구체적인 데이터와 판단기준을 마련하기 위해 과잉진료 규모 등 현황 파악부터 나선다. 과잉진료 고위험군과 정상집단 간 진료현황을 비교해 피해자 속성, 의료기관 이용행태별 평균 진료비, 입원율, 진료기간 등 특성을 상세하게 분석하기로 했다. 이를 바탕으로 과잉진료 고위험군의 특성을 분석한 관리지표를 만들고, 경미 사고 유형별로 부상 위험 등을 측정하는 분석 보고서를 작성할 예정이다. 보험연구원과 공동연구를 진행하기로 했다.

출처 : 헤럴드경제 한희라 기자

