

잡학서고 지식의 기록

통계학의 가장 위대한 정리: 중심극한정리(CLT)

원문 주소

<https://www.gklibrarykor.com/4487/>

PDF 미리보기 · 이 글의 일부 내용입니다.

우리가 살고 있는 현실 세계의 데이터는 결코 정돈되어 있지 않습니다. 어떤 데이터는 한쪽으로 치우쳐 있고, 어떤 데이터는 여러 개의 봉우리를 가지며, 또 어떤 데이터는 아무런 규칙도 없이 불규칙하게 흩어져 있습니다. 그러나 통계학자들은 이처럼 무질서한 데이터 속에서도 전체의 특성을 정확하게 추정해 냅니다. 대한민국 전체 성인의 평균 키나 전국 선거의 지지율을 조사할 때, 모든 사람을 전수조사하지 않고도 높은 정확도로 결과를 예측할 수 있는 비결은 무엇일까요? 그 중심에는 통계학에서 가장 아름답고 강력한 정리로 꼽히는 중심극한정리(Central Limit Theorem, CLT)가 있습니다. 통계학에서 우리가 진짜 알고 싶어 하는 전체 집단을 모집단(Population)이라고 하고, 현실적인 제약으로 인해 모집단에서 일부를 뽑아낸 집단을 표본(Sample)이라고 합니다. 문제는 우리가 분석하려는 모집단이 정규분포(종 모양의 대칭 분포)를 따르는지 사전에 알기 어렵다는 점입니다. 모집단의 분포를 정확히 몰라도 추론할 수 있지만, 사용할 방법에 필요한 표본 추출 방식과 분포의 조건은 확인해야 합니다. 중심극한정리의 기본형은 서로 독립이고 동일한 분포를 따르는 관측