

잡학서고 지식의 기록

마지막 한 장이 가장 어렵다, 쿠폰 수집가 문제(Coupon Collector's Problem)

원문 주소

<https://www.gklibrarykor.com/10099/>

PDF 미리보기 · 이 글의 일부 내용입니다.

과자 봉지마다 들어 있는 스티커를 모아 본 적이 있다면, 처음에는 봉지를 뜯을 때마다 새 스티커가 나와 금방 다 모을 것 같았던 기억이 있을 겁니다. 그런데 앨범이 거의 다 찰 무렵부터는 이미 가진 스티커만 계속 나오고, 마지막 한두 장은 좀처럼 나오지 않습니다. 운이 나빠서가 아니라 확률이 원래 그렇게 움직이기 때문입니다. 이 현상을 수학으로 풀어낸 것이 바로 ‘쿠폰 수집가 문제(Coupon Collector's Problem)’입니다. 쿠폰 수집가 문제란 쉽게 말해 ‘여러 종류 가운데 하나가 무작위로 나올 때, 모든 종류를 적어도 한 번씩 모으려면 평균 몇 번을 뽑아야 하는가’를 묻는 확률 문제를 의미합니다. 답은 생각보다 훨씬 크고, 그 대부분이 마지막 몇 종류를 기다리는 데 쓰인다는 점이 핵심입니다. 본 자료에서는 쿠폰 수집가 문제가 무엇인지부터, 평균 횟수를 계산하는 원리, 숫자로 보는 구체적인 예시, 그리고 이 문제가 실생활과 기술에서 쓰이는 곳과 주의할 점까지 차례로 살펴보겠습니다. 마지막 한 장이 가장 비싸다, 쿠폰 수집가 문제 쿠폰 수집가 문제란 무엇인가 이 문제의 이름은 시리얼 상자마다 쿠폰을 한 장씩 넣고, 모든 종류를 모으면 상품을 주던