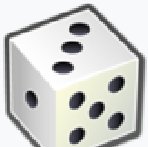
위키백과

우리 모두의 백과사전

확률

통계학 시리즈의 일부

확률론



확률 (공리) · 결정론 (시스템) · 비결정론 · 무작위성

확률공간 · 표본공간 · 사건 (전체 포괄 사건 · 근원사건 · 상호 배타성 · 결과 · 한원소) · 시행 (베르누이 시행) · 확률분포 (베르누이 분포 · 이항분포 · 정규분포) · 확률 측도 · 확률변수 (베르누이 과정 · 연속/이산 · 기댓값 · 마르코프 연쇄 · 관측값 · 무작위 행보 · 확률과정)

여사건 · 결합 확률 · 주변 확률 · 조건부 확률

독립 · 조건부 독립 · 전체 확률의 법칙 · 큰 수의 법칙 · 베이즈 정리 · 불의 부등식

벤 다이어그램 · 트리 다이어그램

v · t · e (https://ko.wikipedia.org/w/index.php?title=%ED%8B%80:%ED%99%95%EB%A5%A0%EB%A1%A0&action=edit)

확정성

불가지론 · 근삿값 · 믿음 · 확정성 · 의심 · 결정론 · 인식론 · 가류주의 · 운명론 · 가설 · 정당화 · 허무주의 · 확률 · 과학 이론 · 회의주의 · 유아론 · 이론 · 진리 · 불확정성

확률(確率, probability)은 어떤 일이 일어날 가능성을 측량하는 단위로 비율이나 빈도로 나타낸다.[1][2]

확률에는 수학적 확률과 경험적 확률이 있다. 수학적 확률은 모든 경우의 수에 대해 그 일이 일어날 확률을 수학적으로 계산한 것이다. 수학적 확률은 모든 경우의 수에 대한 그 일이 일어날 경우의 수의 비로 나타낼 수 있다. 예를 들어 주사위의 각 면은 동일한 기회를 갖고 있기 때문에 일어날 수 있는 모든 경우의 수는 6이고 그 중에 한 면이 위가 될 경우는 1이다. 따라서 주사위의 어떤 눈이 나올 확률은 $\frac{1}{6}$이 된다.[3] 반면 경험적 확률은 실제 그 일을 무수히 반복하였을 때 나타나는 확률로 기존의 경험을 바탕으로 한 추측 값이다.[4]

베이즈 추론은 실제 어떤 일이 무수히 반복되면 경험적 확률과 수학적 확률이 같은 값을 갖는다고 본다. 주사위 던지기나 동전 던지기와 같은 단순한 일이 아니라 사회 현상과 같은 복잡한 일을 다룰 때는 모든 경우의 수를 전부 살펴 볼 수 없기 때문에 표본을 조사하여 결론을 도출하거나 미래를 예측한다. 이런 경우 엄밀한 수학적 확률은 계산할 수 없기 때문에 경험적 확률을 도입할 수 밖에 없다.[5] 베이즈 추론에 따른 경험적 확률은 보험을 비롯한 각종 통계에 도입되고 있다.

수학적 확률은 물리학, 화학, 생물학 등의 과학을 비롯하여 철학이나 도박과 같은 분야에 이르기까지 광범위하게 사용된다. 예를 들어 통계 역학은 다루는 대상이 무수히 많은 경우 이를 확률적으로 계산하고[6] 양자 역학은 물질과 에너지의 상호 작용을 확률로 계산한다.[7]

비율로 표시되는 확률은 0에서 1 사이의 값을 갖는다. 확률 0 은 그 일이 절대로 일어나지 않는 것을 의미하고, 확률 1은 그 일이 확실히 일어난다는 것을 뜻한다.[8][9] 확률은 0.1 과 같은 소수, 10 %와 같은 백분율, $\frac{1}{10}$과 같은 분수 등으로 표현한다. 확률이 0이라는 것은 정해진 경우의 수 가운데 어떤 일이 일어나지 않는 다는 것을 의미하기 때문에 예상 외의 일이 일어날 가능성까지 불가능하다고 보는 것은 아니다. 예를 들어 다트 게임에서 점수는 선으로 구획된 칸을 기준으로 계산되지만, 다트가 표적 칸 사이의 선에 맞을 가능성도 있다. 이런 일이 일어날 경우는 다트의 점수 체계에선 확률 0 이지만, 실제 게임에선 종종 일어난다.[10] 이는 다트 게임 표적의 표본공간이 각 칸의 점

수로만 이루어져 있고 선에 맞을 경우를 고려하고 있지 않기 때문이다. 확률은 집합론으로 정의되며 어떤 일이 일어날 모든 경우의 수를 표본공간으로 하고 그 가운데 어떤 일이 일어날 경우의 수를 표본으로 한다.[11] 예를 든 다트 표적지의 경우 선에 맞을 경우와 빗나갈 경우도 표본공간에 포함시키면 확률 0과 불가능은 같은 의미가 된다.

확률론은 확률을 다루는 수학의 분야이다. 수학적 확률은 수학의 다른 분야와 같이 공리를 바탕으로 연역적으로 추론되며 이를 확률의 공리라고 한다. 확률의 공리는 수학의 다른 분야보다 뒤늦게 정리되었는데 1933년 안드레이 콜모고로프가 《확률론의 기초》에서 정립하였다.[12]

개념과 정의

확률은 모든 일어날 수 있는 경우의 수 가운데 어떤 일이 일어날 가능성을 나타낸다. 고전확률론을 정립한 라플라스는 일어나는 각각의 사건이 독립적일 때 확률은 모든 경우의 수에 대한 어떤 일이 일어날 경우의 수의 비라고 정의하였다. 따라서 확률을 P, 모든 경우의 수를 S, 어떤 일이 일어날 경우의 수를 A라고 하면, A가 일어날 확률 P(A)는 다음과 같이 나타낼 수 있다.[13]

$$P(A) = \frac{A}{S}$$

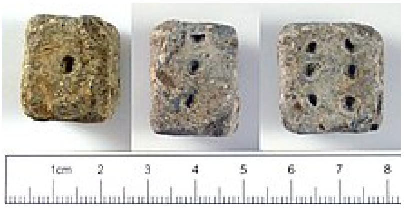
고전적 확률 계산은 실험에서 나타날 일들이 서로 영향을 주지 않아 독립적이며 무작위적이고 모든 경우의 수가 잘 정의 되어 있을 때 가능하다. 즉 주사위 던지기나 포커 게임과 같이 모든 경우의 수를 이미 알고 있고 그 가운데 어떤 일이 일어날 지도 이미 알고 있을 때 계산할 수 있다. 이를 확률의 고전적 정의라고 한다.[14]

확률을 계산할 때, 각각의 경우를 원소로 하는 집합을 생각할 수 있다. 이 경우 전체 집합 S는 모든 경우의 수를 원소로 하는 표본공간이고 그 가운데 어떤 일이 일어날 경우 A, B, C, … 등등은 집합 S의 부분 집합이 된다. 이렇게 생각하면 확률은 표본공간에서 어떤 일이 일어날 경우를 선택하여 표본을 구성하는 함수로 표현될 수 있다. 이를 확률의 공리적 정의라고 한다.[11]

확률을 고전적으로 정의하거나 또는 공리적으로 정의하거나 이 둘은 모두 어떤 일이 일어날 빈도를 계산하는 일이다. 그러기 위해서는 실험 자체가 잘 정의되어 있고 일어날 경우의 수가 분명해야 한다. 그러나 현실 세계에서 사건들은 정의도 불분명하고 전체 경우의 수도 헤아리기 어려운 경우가 많다. 예를 들어 생명보험은 질병이나 사망의 확률을 근거로 보험료를 책정해야 하지만 이런 일들은 독립적이지도 않고 언제 어떻게 발생할 지 예측하기도 힘들다. 빈도를 계산하기 어려운 확률의 책정을 위해 베이즈 확률론이 쓰인다. 베이즈 확률론은 확률을 어떤 일이 일어날 지에 대한 믿음을 근거로 한다. 베이즈 확률론은 논리적 추론을 통해 모든 경우의 수와 일어날 것으로 "믿어지는" 일의 경우의 수를 책정하여 확률을 정한다.[15] 이렇게 정해진 확률은 단지 추론이기 때문에 사전 확률이라 부른다. 한편 일정 기간을 기준으로 실제 일어난 일을 측정하여 확률을 계산할 수 있다. 이를 사후 확률이라고 한다. 확인된 사후 확률은 사전 확률 갱신을 위한 데이터로 사용된다. 베이즈 정리는 데이터가 축적됨에 따라 사전 확률의 정확도를 올릴 수 있음을 보여준다.[16]

역사

이 부분의 본문은 확률의 역사입니다.



고대 로마 말기의 주사위

주사위 게임과 같은 도박은 고대에도 이루어졌다. 고대 로마에서 이루어진 주사위 게임은 결과를 사람의 의지로 어찌할 수 없기 때문에 운명의 여신 포르투나가 개입하는 하는 것으로 믿어졌는데, 이는 무작위에 대한 고대의 관념을 보여준다.[17]

확률은 도박과 같은 분야에서 이미 1천년대 무렵부터 개념이 정리되기 시작하였지만, 이를 수학적으로 접근한 것은 훨씬 후대의 일이다. 도박에서는 게임에서 이길 기회를 여전히 운명의 신이 개입하는 운에 따른 것으로 보았기 때문에 이를 과학적인 관찰 대상으로 여기지는 않았다. 그 때문에 무작위에 놓인 논리적 배경을 탐구하려는 시도는 늦춰질 수 밖에 없었다.[18] 미국의 확률론 학자 리처드 제프리는 "17세기 중반 이전까지 오늘날 '확률'을 뜻하는 '가능한'을 가리키는 낱말 'probable'은 인정할 수 있는 것을 뜻하는 'approvable'을 의미하였다. 행동이나 의견이 '가능하다'는 것은 '타당하다'라는 의미였다."고 설명하였다.[19]

16세기 이탈리아의 지롤라모 카르다노는 본업이 의사였지만 점성술, 철학, 수학 등에도 관심을 보인 박식가였다.[20] 삼차방정식의 일반 해를 구하기도 한 수학자였던 카르다노는 도박에 빠진 도박 중독자이기도 하였다. 그는 《우연의 게임 지침서》(이탈리아어: Liber de ludo aleae)에서 게임의 조건이 참여자 모두에게 공정할 때 일어날 경우의 수를 계산하여 확률의 개념을 수립하였다. 그러나 카르다노의 지침서는 여전히 수학이라기 보다는 도박에 더 중점을 둔 책이었기 때문에 각종 속임수와 같은 것들도 함께 설명하였다.[21]

지롤라모 카르다노

카르다노는 게임의 참가자가 이길 가능성을 따지는 승산을 정의하였다. 게임 참가자가 이길 확률이 p 라면 승산 $Odds$는 아래와 같이 계산된다.

$$Odds = \frac{p}{1-p}$$

위 식에서 분모 $1-p$ 는 모든 경우의 수가 일어날 확률 1에서 승리할 확률 p 를 뺀 게임에 질 확률을 의미한다.[22] 도박은 이기는 것이 목적이기 때문에 승산은 이길 확률과 질 확률의 비율인 승산비로 나타낸다. 승산비는 오늘날 게임 이론에 기초한 연구에서 중요한 도구로 활용된다.[23]

18세기 블레즈 파스칼과 피에르 드 페르마는 수학적 확률 계산의 기초를 정립하였다.[24] 당시 파리의 유명한 도박사였던 앙트완 공보(Antoine Gombaud[* 1])는 포인트라고 불리던 점수 내기 도박을 즐겼다. 어느 날 게임을 도중에 중단하게 된 공보는 파스칼에게 게임에 걸린 판돈을 공정하게 분배할 방법을 물었다. 파스칼은 문제의 해답을 구하면서 마랭 메르센을 통하여 페르마와 의견을 주고 받았다. 둘은 서로 다른 방법으로 공보의 문제에 대한 동일한 결론을 도출하였고 이는 모든 확률 문제를 해결할 수 있는 일반적 방법으로 발전하였다.[25]:65-66

포인트는 판돈을 걸고 게임을 하여 특정한 점수에 먼저 도달한 사람이 이기는 주사위 게임이었다. 두 참가자 A와 B가 3점을 먼저 내는 사람이 이기는 것으로 하고 게임을 하다가 A는 2점 B는 1점을 낸 상태에서 게임을 중단하였다. 둘 모두가 앞으로 점수를 얻을 확률이 동일할 때 판돈은 아래와 같이 나눌 수 있다.[26]

A는 한 번 만 더 이기면 게임에서 승리하고, B는 두 번을 연달아 이겨야 하므로 게임은 최대 2 번까지 더 할 수 있다. 따라서 두 변수의 제곱에 대한 이항정리로 나타내면,

$$(A+B)^2 = 1 \cdot A^2 + 2 \cdot AB + 1 \cdot B^2$$

위 식의 각 항 가운데 $1 \cdot A^2$, $2 \cdot AB$는 A의 승리를 뜻하고, B가 연거푸 이기는 경우는 $1 \cdot B^2$ 밖에 없기 때문에 모든 항의 계수를 더한 4를 분모로 하고 각자가 이길 경우를 분자로 한 확률에 따라 A가 전체 판돈의 $3/4$을, B가 $1/4$을 나누어 가지면 된다.

야코프 베르누이는 《추론술》(라틴어: Ars Conjectandi, 1713년)을 통하여 확률론을 수학의 한 분야로 자리잡게 하였다.[27] 아브라암 드무아브르는 《확률론》(영어: The doctrine of chances, 초판 1718, 2판 1738년)을 통하여 확률론의 수학적 기초를 세웠고 확률분포를 설명하였다. 특히 2판에서 정규분포를 언급하였다.[28] 정규분포는 훗날 카를 프리드리히 가우스에 의해 엄밀하게 정의 되었다.[29] 라플라스는 집합론을 도입하여 확률을 재정의 하였다. 또한 조건부 확률, 주변 분포와 같은 비독립적 사건을 포함하여 확률론을 일반화 하였다.[30]

아돌프 케틀레는 자연 과학에 이용되던 확률 개념을 사회 과학으로 확장하였다. 케틀레는 1835년 《인간과 그 능력의 개발에 관한 논고: 사회물리학 시론》에서 사회 현상 역시 확률과 통계를 적용할 수 있다고 보았고 사회 현상을 대표할 수 있는 "평균적 인간"이라는 개념을 도입하였다.[31]

토머스 베이즈는 모든 경우의 수의 집합인 S가 불명확할 때 확률을 계산하는 베이즈 정리를 통해 베이즈 확률론의 기초를 만들었다. 그는 선험적인 사전확률과 일정 기간을 통해 측정한 사후확률의 되먹임을 통해 여러 사회 현상에 유용한 확률 계산법을 고안하였다.[32]

베이즈 확률론은 조건부 사건이 연속하여 발생하는 것을 전제로 한다. 그런데 앞서 일어난 일이 뒤에 일어날 일과 아무런 관련이 없는 경우도 있기 마련이다. 《한비자》 오두편의 수주대토(守株待兎)는 우연한 일이 반복되길 기다리는 어리석음을 비웃는 말이다.[* 2] 1906년 안드레이 마르코프는 과거의 사건이 현재에는 조건부로 작용하지만 미래에는 영향을 주지 않는 마르코프 확률 과정을 정의하여 이러한 경우의 확률을 계산할 수 있도록 하였다.[33]

주요 요소


포커는 확률 분포의 차이를 이용한 게임이다.

일어날 수 있는 모든 경우의 수 가운데 선택된 각각의 경우를 사건이라고 하며, 더 이상 분해 될 수 없는 기본 사건과 이들이 결합한 복합 사건으로 구분된다. 또한 기본 사건은 그 성격에 따라 독립 사건 과 종속 사건, 상호 배타 사건 등으로 구분된다.

현대 수학에서 확률은 집합론과 함수에 의해 설명된다. 즉 확률은 어떤 일이 일어날 모든 경우의 수의 집합을 정의역인 표본 공간으로 하고 그것이 일어날 확률의 집합인 확률 공간을 치역으로 하며 함수인 확률 변수에 의해 정의역에서 치역으로 사상된다.

하나의 정육면체 주사위 던지기와 같은 기본 사건은 각각의 경우의 수가 나타날 확률이 균등하지만, 둘 이상이 되면 나타나 수 있는 확률의 분포가 동일하지 않다. 복합 사건에서 어떤 것은 보다 많이 나타날 가능성이 있고 어떤 것의 확률은 희박하다. 포커는 이러한 확률의 차이를 바탕에 둔 카드 게임이다. 확률 분포의 정도는 확률 질량 함수나 확률 밀도 함수를 통해 계산할 수 있으며 이렇게 계산된 정도에 따라 기대값을 생각할 수 있다.

사건의 종류

독립 (확률론) 문서를 참고하십시오.

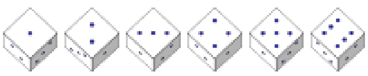
일어나는 사건들이 서로 아무런 영향을 주지 않을 때 독립 사건이라고 한다.[34] 반면에 둘 이상의 사건이 서로 영향을 미친다면 종속 사건이라고 한다. 정육면체 주사위 던지기에서 나오는 눈의 값은 서로 영향을 미치지 않기 때문에 독립 사건이지만, 흰 공과 검은 공이 섞여 있는 주머니에서 공을 꺼내는 사건의 확률은 종속적이다. 앞서 꺼낸 공 때문에 확률이 변하기 때문이다. 예를 들어 흰 공 5 개와 검은 공 5 개가 들어 있는 주머니에서 흰 공을 꺼낼 확률은 50% 이지만 처음 꺼낸 공이 흰 공이었을 때 두 번째 공도 흰 공일 확률은 $^{4}/_{9}$로 변하게 된다. 이미 흰 공 하나가 밖으로 나와 있기 때문이다.

배반 사건 또는 배타 사건은 동시에 두 가지가 일어날 수 없는 경우를 가리킨다. 정육면체 주사위를 던졌을 때 홀수이면서 동시에 짝수일 수는 없기 때문에 두 사건은 서로 배타적이다.[34] 반면에 두 사건이 동시에 고려되어야 한다면 결합 사건이라고 한다.[35]

표본 공간과 확률 공간

표본 공간 및 확률 공간 문서를 참고하십시오.

어떤 일이 일어날 확률은 일어날 수 있는 모든 경우의 수에 대한 선택된 일이 일어날 경우의 수의 비율로 정의할 수 있다. 예를 들어 정육면체 주사위를 한 번 던져 나올 수 있는 모든 경우의 수는 6이다.



이 가운데 짝수가 나올 경우는 2, 4, 6의 세 가지이므로, 정육면체 주사위를 한 번 던져 짝수가 나올 확률은 $^{3}/_{6}$ 이고 약분 하면 $^{1}/_{2}$이 된다. 0.5 또는 50 % 로 표기할 수도 있다.

이 때 모든 경우의 수를 원소로 하는 집합을 표본 공간이라고 하고 일반적으로 Ω 또는 S 로 표기한다.[36] 한편 위의 예에서 짝수가 나오는 경우처럼 선택된 일이 일어나는 경우를 원소로 하는 집합을 사건이라고 하고 일반적으로 F 로 표기한다. 따라서 사건 F는 언제나 표본 공간 S의 부분 집합이다.[36]

$$F \subseteq S$$

확률 공간은 표본 공간에서 사건이 일어날 범위를 측정하는 측도 공간이다. 사건 F가 아무런 원소도 갖지 않는 공집합이라면 확률은 0 이 되고 사건 F에 속하는 원소가 전체 집합인 표본 공간 S와 같다면 확률은 1이 된다.

정육면체 주사위를 한 번 던졌을 때 나오는 눈을 원소로 하는 표본 공간 S는 다음과 같다.

$$S = \{1, 2, 3, 4, 5, 6\}$$

이때 짝수가 나오는 사건 F는 다음과 같은 부분 집합이다.

$$F = \{2, 4, 6\}$$

표본 집합 S에 대한 사건 F의 확률은 다음과 같이 표현된다.

$$P(F) = \frac{F}{S} = \frac{1}{2}$$

한편 정육면체 주사위의 경우 짝수가 나오지 않으면 홀수이기 때문에 이 경우는 사건 F의 여 집합으로 생각할 수 있다.

확률 변수

이 부분의 본문은 확률 변수입니다.

현대 수학에서 확률은 표본 공간에서 확률 공간으로 사상되는 함수로 취급된다. 표본 공간의 원소인 각각의 경우의 수를 일정한 확률로 계산할 수 있는 것은 확률 변수의 작용 때문이다. 표본 공간은 확률 함수의 정의 역으로 동전 던지기처럼 단 두 가지 경우를 원소로 할 수도 있고 실수 전체를 원소로 할 수도 있다. 그러나 확률 변수에 의해 사상 된 확률의 치역은 0과 1사이의 실수 값을 갖는다.[37]

확률 분포

이 부분의 본문은 확률 분포입니다.

주사위 두 개를 한 번에 던져 나오는 눈의 수를 더한다고 해 보자. 두 눈의 합이 6이 될 확률은 얼마나 될까? 주사위 두 개를 한 번에 던졌을 때 나올 수 있는 모든 경우의 수는 36가지로 이 때 표본 공간 S는 다음과 같은 순서쌍을 원소로 하는 집합이다.

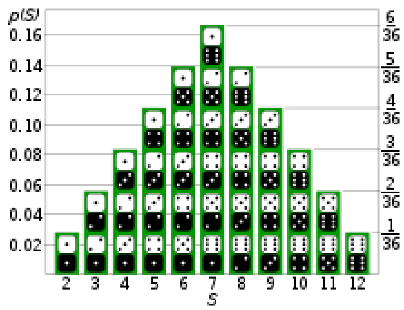
$$S = \{(1,1), (1,2), (1,3) \cdots\cdots (6,4), (6,5), (6,6)\}$$

이 가운데 더하여 6 이 되는 순서쌍을 원소로 하는 집합 사건 F는 다음과 같다.

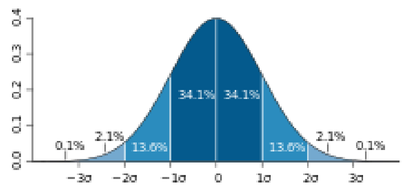
$$F = \{(1,5), (2,4), (3,3), (4,2), (5,1)\}$$

따라서 두 눈의 합이 6이 될 확률 $P(F)$는 $^5/_{36}$로 약 13.89 %이다. 반면에 더하여 2가 되는 경우는 두 눈 모두 1 인 경우 밖에 없기 때문에 확률은 $^1/_{36}$로 약 0.28 %에 불과하다.

이와 같이 서로 다른 확률을 갖는 사건들의 분포를 확률 분포라고 한다. 확률 분포는 주사위 던지기나 동전 던지기와 같이 불연속적인 값일 수도 있고, 신입생의 키나 몸무게와 같이 연속적으로 늘어선 구간의 어느 값으로 나타날 수도 있다. 불연속적인 값을 갖는 확률 분포를 이산 확률 분포라고 하고 연속적인 경우를 연속 확률 분포라고 한다. 정규분포는 가장 널리 쓰이는 연속 확률 분포이다.



주사위 둘을 던졌을 때 나오는 눈의 합이 보여주는 이산 확률 분포



정규분포 곡선

확률 분포 함수

확률 질량 함수 및 확률 밀도 함수 문서를 참고하십시오.

확률의 분포가 확률 변수에 따라 차이가 나는 점을 생각할 때 확률의 분포 자체를 표현하는 확률 분포 함수를 생각할 수 있다. 연속적이지 않은 이산형 확률 변수 가 취할 수 있는 값이 확률 질량 함수이다. 반면에 연속적인 확률 변수의 경우 일정 구간의 확률을 넓이로 구할 수 있기 때문에 확률 밀도 함수라고 한다. 질량이나 밀도라는 명칭은 두 함수의 계산 법이 물리적 질량이나 밀도를 구하는 식과 같기 때문에 붙인 것이다.[38] 빗물이 무작위로 쟁반 위에 떨어진다고 하자, 쟁반의 모양이 완전한 원반이라면 원의 중심에서 일정 반지름까지 빗방울이 떨어질 확률은 전체 원반 넓이에 대한 특정 반지름까지 원의 넓이 비율로 계산 될 것이다. 확률 밀도 함수는 이와 같이 표본 공간이 무한이 많은 경우로 이루어 질 때 확률의 분포를 표현하기 위해 쓰인다. 단, 확률 분포 함수는 확률 변수의 분포 만을 다룰 뿐 확률 자체를 나타내지는 않는다.

기댓값

🔍 이 부분의 본문은 기댓값입니다.

확률 분포 함수에 따라 어떤 사건이 가장 많이 일어나는 지 알 수 있다. 기댓값은 확률 분포의 대표적인 중심 값으로 확률분포의 성격을 결정짓는 확률적 평균치이다. 기댓값은 산술적 평균이 아니며 확률의 예측에 따른 가중 평균이다.[39]

예를 들어 정육면체 주사위 두 개를 동시에 던졌을 때 나올 수 있는 확률 변수 X에 대한 기댓값 E(X)는 다음과 같이 계산된다.[40]

$$E(X) = \Sigma x_i P(X = x_i)$$

x_i는 X의 i 번째 원소, $P(x_i)$는 그 원소의 확률

$$\begin{aligned}
&= x_1p_1 + x_2p_2 + x_3p_3 + \cdots\cdots + x_{10}p_{10} + x_{11}p_{11} \\
&= 2\times\frac{1}{36} + 3\times\frac{2}{36} + 4\times\frac{3}{36} + 5\times\frac{4}{36} + 6\times\frac{5}{36} + \\
&\quad 7\times\frac{6}{36} + 8\times\frac{5}{36} + 9\times\frac{4}{36} + 10\times\frac{3}{36} + 11\times\frac{2}{36} + 12\times\frac{1}{36} \\
&= \frac{2+6+12+20+30+42+40+36+30+22+12}{36} = \frac{252}{36} = 7
\end{aligned}$$

확률의 합과 곱

다시 한 번 주사위 던지기를 생각해 보자. 정육면체 주사위를 한 번 던졌을 때 나오는 값이 2의 배수일 사건을 F2 라 하고, 3의 배수일 사건을 F3이라고 하면 두 사건은 다음과 같이 나타낼 수 있다.

$$F2 = \{2, 4, 6\}$$
$$F3 = \{3, 6\}$$

이제 정육면체 주사위를 한 번 던져 2의 배수 이거나 3의 배수일 경우의 확률은 두 사건 F2와 F3의 합집합 $F2 \cup F3$로 생각할 수 있다. 이 때 2의 배수이기도 하고 3의 배수이기도 한 6은 중복 되는 교집합이므로 한 번 빼줘야 한다.

$$F2 \cup F3 = F2 + F3 - F2 \cap F3 = \{2, 4, 6\} + \{3, 6\} - \{6\} = \{2, 3, 4, 6\}$$

따라서 정육면체 주사위를 한 번 던져 2의 배수 이거나 3의 배수일 경우의 확률은 다음과 같다.[41]

$$P(F2 \cup F3) = \frac{4}{6} = \frac{2}{3}$$

한편 정육면체 주사위를 두 번 던져 처음 나온 값은 2의 배수, 두 번째의 것은 3의 배수일 경우를 생각해 보면, 다음과 같이 두 확률의 곱으로 나타낼 수 있다. 두 사건이 동시에 일어나는 경우를 곱사건이라고 한다.

$$P(F2 \text{ and } F3) = P(F2 \cap F3) = P(F2)P(F3)$$

이를 계산하면,

$$P(F2) \times P(F3) = \frac{3}{6} \times \frac{2}{6} = \frac{6}{36} = \frac{1}{6}\ .$$

생방송 행복드림 로또 6/45는 1에서 45개의 숫자 가운데 6 개를 맞추는 곱사건이고, 한 숫자를 맞출 때마다 골라야 하는 숫자와 고를 수 있는 숫자가 하나씩 줄어드는 종속 사건이기 때문에 6 개의 숫자 모두를 맞출 확률은 다음과 같이 여섯 가지 확률을 곱셈하여 계산할 수 있다.

$$\frac{6!}{45^{\underline{6}}} = \frac{6}{45} \times \frac{5}{44} \times \frac{4}{43} \times \frac{3}{42} \times \frac{2}{41} \times \frac{1}{40}$$
$$= \frac{720}{5,864,443,200} = \frac{1}{8,145,060}$$

조건부 확률

이 부분의 본문은 조건부 확률입니다.

조건부 확률은 주어진 사건이 일어났다는 가정 하에 다른 한 사건이 일어날 확률을 뜻한다.[42] 사건 A가 일어났을 때 사건 B가 일어날 조건부 확률은 $P(B|A)$로 나타내고 A가 일어날 확률 $P(A)$를 분모로 하고 두 사건이 동시에 일어날 곱사건의 확률 $P(A \cap B)$를 분자로 하여 나타낼 수 있다.

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

조건부 확률의 널리 알려진 예시로는 몬티 홀 문제가 있다. 세 개의 문 가운데 하나에 경품이 있다. 참가자가 문 하나를 선택하면 진행자는 남은 두 문 가운데 경품이 없는 문을 하나 열어 보여 주고는 참가자에게 선택을 바꿀 것인지 묻는다. 참가자 입장에선 어떤 것이 유리한가? 이 문제는 미국의 텔레비전 쇼 《거래를 합시다(Let's Make a Deal)》에서 유래한 것이다. 대부분의 참가자는 자신의 선택을 유지하였는데, 세 문 가운데 하나에 경품이 없다는 것을 확인하였기 때문에 당첨 확률이 $^1⁄_3$에서 $^1⁄_2$로 올라갔다고 생각하기 때문이다. 그러나 선택을 바꾸면 당첨 확률은 $^2⁄_3$으로 올라간다.[43]

같이 보기

- 게임 이론
- 정보 이론
- 측도
- 확률 과정
- 통계
- 확률론
- 베이즈 통계학

각주

내용주

1. 공보는 귀족이 아니었지만 당시 "메르의 기사"라 불렸다. 이 때문에 종종 "드 메르", "드 메레"로 불리기도 한다.
2. 수주대토(守株待兎)는 어쩌다 토끼가 나무에 부딛쳐 죽는 걸 본 사람이 그 나무 옆에서 다음 토끼가 부딛히기를 기다린다는 의미의 고사성어이다.

참조주

1. 확률, 확률 공리 (http://www.ktword.co.kr/test/view/view.php?m_temp1=1644), 정보통신용어해설
2. 내일 비올 확률 50% … 그 뜻은? 숫자가 정보다. 확률을 이해하라 (https://dbr.donga.com/article/view/1206/article_no/6715/ac/magazine), 동아비지니스리뷰, 2014년 11월
3. 수학적 확률 (https://www.scienceall.com/%ec%88%98%ed%95%99%ec%a0%81-%ed%99%95%eb%a5%a0mathematical-probability/?term_slug=), 사이언스올
4. 경험적 확률 (https://www.scienceall.com/%EA%B2%BD%ED%97%98%EC%A0%81-%ED%99%95%EB%A5%A0empirical-probability-%E7%B6%93%E9%A9%97%E7%9A%84-%E7%A2%BA%EF%A5%A1-2/),사이언스올
5. 강승택 이상은,〈표본조사에서 베이지안적 추정방법 (http://kostat.go.kr/file_total/eduSri/22-01-02.pdf)〉, 《통계연구》, 제22권 제1호, 2017년
6. 통계의 날 살펴 본 과학 속 통계 (https://scienceon.kisti.re.kr/srch/selectPORSrchTrend.do?cn=SCTM00087014), 사이언스온, 2010년 9월 1일
7. 양자역학에 확률의 개념을 도입하다 (https://www.sciencetimes.co.kr/news/%EC%96%91%EC%9E%90%EC%97%AD%ED%95%99%EC%97%90-%ED%99%95%EB%A5%A0%EC%9D%98-%EA%B0%9C%EB%85%90%EC%9D%84-%EB%8F%84%EC%9E%85%ED%95%98%EB%8B%A4/), 사이언스타임스, 2020년 2월 5일
8. "Kendall's Advanced Theory of Statistics, Volume 1: Distribution Theory", Alan Stuart and Keith Ord, 6th Ed, (2009), ISBN 978-0-534-24312-8.
9. William Feller, *An Introduction to Probability Theory and Its Applications*, (Vol 1), 3rd Ed, (1968), Wiley, ISBN 0-471-25708-7.
10. ‘확률이 0 이다’와 ‘불가능하다’의 차이는? (http://pub.chosun.com/client/article/viw.asp?cate=C03&nNewsNumb=20170323791), 정경조선, 2017년 3월 8일
11. 확률의 수학적 정의와 의미 (https://datascienceschool.net/02%20mathematics/06.02%20%ED%99%95%EB%A5%A0%EC%9D%98%20%EC%88%98%ED%95%99%EC%A0%81%20%EC%A0%95%EC%9D%98%EC%99%80%20%EC%9D%98%EB%AF%B8.html), 데이터 사이언스 스쿨
12. Kolmogorov, Andrey (1950) [1933]. 《Foundations of the theory of probability》 (https://archive.org/details/foundationsofthe00kolm). New York, USA: Chelsea Publishing Company.
13. 확률(probability)이란? (http://bigdata.dongguk.ac.kr/lectures/med_stat/_book/%ED%99%95%EB%A5%A0probability%EC%9D%B4%EB%9E%80.html), 의학통계강의, 김진석
14. Jaynes, E. T., 2003, Probability Theory: the Logic of Science, Cambridge University Press, see pg. xx of Preface and pg. 43.
15. ‘빈도’의 통계적 확률과 ‘믿음’의 논리적 확률 (http://times.postech.ac.kr/news/articleView.html?idxno=3053), 포항공대신문
16. Priori Probability, Posteriori Probability 사전 확률, 선험적 확률, 사후 확률, 후험적 확률 (http://www.ktword.co.kr/test/view/view.php?m_temp1=4215), 정보통신용어해설
17. 고대의 주사위는 왜 이 모양이지? (http://www.astronomer.rocks/news/articleView.html?idxno=89251), 이웃집과학자, 2020년 8월 26일
18. John E. Freund, (1973) *Introduction to Probability*. Dickenson ISBN 978-0-8221-0078-2 (p. 1)
19. Jeffrey, R.C., *Probability and the Art of Judgment,* Cambridge University Press. (1992). pp. 54–55 . ISBN 0-521-39459-7
20. 이 주일의 역사 - 수학자 카르다노 사망(1576.9.21) (http://www.busan.com/view/busan/view.php?code=20100920000069), 부산일보, 2010년 9월 20일
21. 확률을 최초로 연구하기 시작한 수학자 ‘카르다노' (http://bc.kyobobook.co.kr/front/subscribe/detailCotents.ink?contents_no=9713&orderClick=7mo), 교보문고, 2020년 8월 14일

22. “*Some laws and problems in classical probability and how Cardano anticipated them* Gorrochum, P. *Chance* magazine 2012” (http://www.columbia.edu/~pg2113/index_files/Gorroochurn-Some%20Laws.pdf) (PDF).
23. 조성경 외, 〈두 모집단 비율 비교에서 상대위험도와 승산비에 기초한 표본크기 추정 (https://e-jhis.org/upload/pdf/21400502.pdf)〉, 《한국보건정보통계학회지》, 제36권 제1호, 2011년
24. Abrams, William, 《A Brief History of Probability》 (https://web.archive.org/web/20170724052656/http://www.secondmoment.org/articles/probability.php), Second Moment, 2017년 7월 24일에 원본 문서 (http://www.secondmoment.org/articles/probability.php)에서 보존된 문서, 2008년 5월 23일에 확인함
25. 사이먼 싱, 박병철 옮김, 《페르마의 마지막 정리》, 영림카디널, ISBN 978-89-8505-597-0
26. 도박사와 수학자 (https://www.sciencetimes.co.kr/news/%EB%8F%84%EB%B0%95%EC%82%AC%EC%99%80-%EC%88%98%ED%95%99%EC%9E%90/), 사이언스타임스, 2004년 7월 19일
27. 야콥 베르누이 (https://www.scienceall.com/%EC%95%BC%EC%BD%A5-%EB%B2%A0%EB%A5%B4%EB%88%84%EC%9D%B4jakob-bernoulli/), 사이언스올
28. 정규분포 (http://www.allthatmath.com/bbs/board.php?bo_table=sub_113&wr_id=304), 올댓매스
29. 정규 분포, 가우시안 분포, 가우스 분포, 가우시안 확률분포 (http://www.ktword.co.kr/word/abbr_view.php?m_temp1=1767), 정보통신용어해설
30. 서적소개 – 확률에 대한 철학적 시론 (피에르 시몽 라플라스 / 지식을만드는지식 / 2009.4.15) (http://chedulife.com.au/%EC%84%9C%EC%A0%81%EC%86%8C%EA%B0%9C-%ED%99%95%EB%A5%A0%EC%97%90-%EB%8C%80%ED%95%9C-%EC%B2%A0%ED%95%99%EC%A0%81-%EC%8B%9C%EB%A1%A0-%ED%94%BC%EC%97%90%EB%A5%B4-%EC%8B%9C%EB%AA%BD-%EB%9D%BC%ED%94%8C/), 크리스천에듀라이프
31. 과학과 수학의 불확실성 – (5) 19세기의 과학과 수학 (https://greenacademy.re.kr/archives/5032), 녹색아카데미
32. 베이즈 정리 (https://datascienceschool.net/02%20mathematics/06.06%20%EB%B2%A0%EC%9D%B4%EC%A6%88%20%EC%A0%95%EB%A6%AC.html), 데이터 사이언스 스쿨
33. Markov Process, Markov Chain 마르코프 프로세스, 마르코프 과정, 마코브 과정, 마르코프 모델, 마르코프 연쇄 (http://www.ktword.co.kr/test/view/view.php?m_temp1=4312), 정보통신용어해설
34. 독립 사건, 독립 시행, 통계적 독립, 통계적으로 독립, 종속 사건, 상호 배타적, 배반 사건 (http://www.ktword.co.kr/test/view/view.php?m_temp1=4419&id=1000), 정보통신용어해설
35. 결합 확률, 결합 확률분포, 결합 확률함수, 결합 모멘트, 결합 통계량 (http://www.ktword.co.kr/test/view/view.php?m_temp1=4062&id=1000), 정보통신용어해설
36. Probability Space 확률 공간, 확률적 공간 (http://www.ktword.co.kr/test/view/view.php?m_temp1=1662), 정보통신용어해설
37. 확률 변수와 확률 분포 (http://kanggc.iptime.org/stat/chap5/chap5.pdf), e러닝 지원센터
38. 확률분포함수 (https://datascienceschool.net/02%20mathematics/06.04%20%ED%99%95%EB%A5%A0%EB%B6%84%ED%8F%AC%ED%95%A8%EC%88%98.html), 데이터 사이언스 스쿨
39. 기대치, 기대값 (http://www.ktword.co.kr/test/view/view.php?nav=2&no=1638&sh=%EA%B8%B0%EB%8C%80%EA%B0%92), 정보통신용어해설
40. 11. 확률 (http://contents.kocw.or.kr/KOCW/document/2015/shinhan/kimeuhee/11.pdf), 이산수학
41. 확률의 계산 (https://mathbang.net/114), 수학방
42. 조건부 확률 (http://www.ktword.co.kr/test/view/view.php?nav=2&no=3731), 정보통신용어해설
43. Monty Hall Problem (https://mathworld.wolfram.com/MontyHallProblem.html), Wofram Mathworld