한국수학사학회지 제23권 제3호 (2010년 8월) 107–119

기댓값에 대한 역사적 고찰

대전송촌고등학교 이종학
mathro@hanmail.net

확률 개념과 지식이 정립되는 단계에서 중요한 의의를 가졌던 기댓값이 확률의 역사에서 초기에 역사적으로 어떻게 다루어졌는가를 알아보는 것은 의미 있는 일이다. 이에 본 연구에서는 기댓값의 개념과 정의가 변화하고 정선되어 온 역사적 과정에서 이 발전에 기여한 몇 가지 문제와 함께 그 문제의 해결 과정을 알아봄으로써 확률에서 기댓값의 의미를 다시 한번 고찰해 보고자 한다.

주제어： 확률, 기댓값, 공평한 분배.

1 서론

교과서에 나오는 수많은 수학적 지식과 개념들은 오랜 기간을 통해 수학자들의 발견과 논쟁을 통해 변화되고 발전되면서 정립되어 온 것이다. 하지만 학생들에게 수학적 지식과 개념들이 처음 제시될 때는, 그 정선되고 세련되어진 지식이나 개념들이 왜 나타나게 되었는지에 대한 역사적 상황이나 배경을 보여주지는 않는다. 또한 그 지식과 개념들이 어떻게 변화되고 정선되었으며 세련되어 갔는지에 대한 과정은 뒤로 감춰지고, 많은 수학적 지식들이 처음부터 정교하게 나타난 것과 같은 모습으로 학생들에게 제공되어 진다.

따라서 교과서를 통해 수학을 처음 배우는 학생들은 그 수학적 지식이 태동하게 된 이유와 역사적 배경이 무엇인지, 수학적 지식이 전개되는 과정에서 어떠한 어려움이 있었으며, 그 어려움을 해결하는 과정에서의 논쟁거리는 무엇이었는지, 또 그 논쟁거리가 어떤 과정을 거쳐서 어떻게 해결되었고, 그 과정에서 어떤 수학적 사실이 뒤쳐지고 그에 따라 또 다른 어떤 수학적 사실이 보강되게 됐는지 등 수학적 지식이 발생한 생생한 역사적 상황에 대해서는 거의 알 수가 없다.

물론 역사적으로 오랜 시간을 거쳐 이루어진 수학 지식 발생의 모든 과정을 교실 수업 내에서 그대로 재현시켜야 할 필요는 없다. 그렇지만, 학생의 수학적 개념을 형성하는데

도움을 줄 수 있는 역사적 상황을 선택하고, 이 상황들을 효과적으로 연관을 짓는 교수학적 변환 작업을 통해 학생들이 교실 상황에서 수학자로써 수학하는 활동을 경험해 볼 수 있도록 하는 것은 학교 수학이 지녀야 할 또 다른 모습인 발견적이고 형성적인 측면에서 볼 때 수학을 공부하는 학생들에게 어느 정도는 의미가 있는 활동일 것이다.

다른 수학 지식과 마찬가지로 확률 영역도 변화하고 발전하면서 하나의 학문 영역으로 정립되어 왔다. 확률론은 20세기 초까지 라플라스(Laplace:1749～1827)의 고전적 확률의 정의를 사용하였지만, 이 정의는 일어날 가능성, 즉 결과로 나올 가능성이 모두 같은 것으로 간주하는 표본공간에서 정의된 확률이기 때문에 조건을 만족하지 않는 상황에서는 확률을 계산할 수가 없다.

예를 들어 바구니(가)에는 검은 공(B) 2개와 흰 공(W) 1개가 들어 있고, 바구니(나)에는 검은 공(B) 1개와 흰 공(W) 1개가 들어 있다. 동전을 던져 앞면이 나오면 바구니(가)에서 공 한 개를 꺼내고, 뒷면이 나오면 바구니(나)에서 공 한 개를 꺼낼 때, 흰 공이 나올 확률을 구하는 문제를 풀 때, 일반적인 풀이인 확률의 합·곱의 법칙을 사용하여 확률을 구하지 않고 고전적 확률의 정의를 사용한다면 다음과 같은 직관적 모순에 빠질 수도 있다.

라플라스의 고전적 정의를 사용하여 확률을 계산할 때, 확률의 분모에 해당하는 모든 경우의 수는 문제 상황에서 직관적으로 다음과 같이 구할 수 있다. 동전을 던져서 앞면이 나오면 검은 공(B) 2개와 흰 공(W) 1개가 들어 있는 바구니 (가)에서 공 한 개를 꺼내고, 뒷면이 나오면 검은 공(B) 1개와 흰 공(W) 1개가 들어 있는 바구니 (나)에서 공 한 개를 꺼내는 시행이다. 가능한 모든 경우의 수의 집합을 S라 하면,

$$S = \{(\text{앞면}, \mathrm{B}), (\text{앞면}, \mathrm{B}), (\text{앞면}, \mathrm{W}), (\text{뒷면}, \mathrm{B}), (\text{뒷면}, \mathrm{W})\}$$

이고, $n(S) = 5$이다. 이때 흰 공이 나올 경우의 수의 집합을 A라 하면,

$$A = \{(\text{앞면}, \mathrm{W}), (\text{뒷면}, \mathrm{W})\}$$

이고, $n(A) = 2$이다. 따라서, 흰 공이 나올 확률 $P(A) = \frac{n(A)}{n(S)} = \frac{2}{5}$ 이다.

하지만 이 풀이는 잘못된 풀이이고, 올바른 해답은 다음과 같다. 바구니(가)에서 흰 공을 꺼낼 확률은 $\frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$이고, 바구니(나)에서 흰 공을 꺼낼 확률은 $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$이다. 따라서, 흰 공이 나올 확률은 $\frac{1}{6} + \frac{1}{4} = \frac{5}{12}$이다.

이처럼 제한적이고 모호했던 확률 영역의 지식들이 공리적 방법에 의한 새로운 논리 구

조로 확립되기까지는 확률 연구의 직접적인 동기가 되는 17세기 초 페르마(Fermat:1601 ~1665)와 파스칼(Pascal: 1623~1662)의 연구에서 상당 부분을 차지하는 기대 금액, 다시 말해 기댓값에 관한 해법으로부터 1933년 콜모고로프(Kolmogorov: 1903~1987)의 공리적 확률의 정의에 이르기까지 400년 이상이 걸렸다. 이러한 확률의 역사에서 중요한 의의를 지녔던 내용인 기댓값이 초기에 어떻게 다루어졌는가를 집중적으로 고찰해 보는 것만으로도 의미 있는 작업일 수 있을 것이다.

학생들이 수학을 공부면서 갖게 되는 의문점 가운데 하나가 "도대체 누가, 왜, 어떻게, 이러한 것을 생각하였을까?"라는 것이다. 이와 같은 의문에 답을 얻지 못하고, 변화하고 발전하는 수학이 아닌 완성된 형태로서의 수학만을 보는 것은 수학의 참다운 모습을 알지 못하는 것일 수 있다. 이에 본 연구에서는 기댓값의 개념과 정의가 변화하고 정립되어 온 역사에 대한 고찰을 통해 기댓값의 의미를 다시 고찰해 보고자 한다.

2 기댓값

2.1 공평한 분배

중세시대에 행해졌던 내기도박은 확률론에 관한 연구를 가져오는 계기가 되었고, 17세기 초 도박판의 상금을 분배하는 분배문제의 해결을 위해 Fermat와 Pascal 사이에 주고받은 문서에 나타난 기대 금액, 즉, 기댓값이 확률론 연구의 직접적인 동기가 되었다.

1500년 경 이탈리아 수학자들은 어떤 예측이 가능한 상황에서 수학적으로 공평한 분배의 방법을 찾기 위해 노력했지만 그 당시에는 확률 이론에 대해 자세히 알지 못했기 때문에 공평한 분배에 대한 올바른 해법을 제시하지는 못하였다.

> 「n번의 게임을 이겨야 전체 내기 돈 1을 차지할 때, A가 더 이겨야 할 게임 수는 a, B가 더 이겨야 할 게임 수는 b인 상황에서 게임이 중단되었다고 하자. 이 때, 내기 돈을 얼마씩 나누어 가져야 하는가?」

이 문제의 해법에 대해서 최초로 언급한 수학자 파치올리(Pacioli:1445~1509)는 $n=6, n-a=5, n-b=2$인 특별한 경우에서 상금을 $5:2$로 나누어 가져야한다는 해법을 제시하였다. 이 해법을 제시한 이유로 그는 이 게임이 최대로 진행 가능한 횟수를 $2(n-1)+1$이라고 하면서, 이 때 내기 돈의 분배는 $\frac{n-a}{2n-1} : \frac{n-b}{2n-1}$이라고 주장했다. 즉, 전에 이긴 게임의 수를 진행이 가능한 게임의 최대수로 나누는 방법을 제안한 것인데, 그의 풀이는 논리적으로 타당하지 못한 해법이었다.

이 후 프랑스의 귀족 드메레(de Méré)는 Pascal에게 우연 게임에 대한 2개의 문제를 보냈고, Pascal은 이 문제들을 해결하고자 노력하였다. 그 당시에는 과학적인 학술지가 없었기 때문에 새로운 결과를 알리기 위해 동료들에게 편지를 보내는 것이 관례였다. Pascal은 그의 해법에 대한 Fermat의 조언을 구했고, 그 관계가 계속 이어지면서 공평한 분배문제의 해법은 수학적으로 체계화될 수 있었다. de Méré의 편지에 있었던 첫 번째 문제는 다음과 같다.

「2개의 주사위를 던져서 적어도 한번 이상 $(6,6)$이 나올 가능성이 $\frac{1}{2}$보다 크려면 최소한 몇 번을 던져야 하는가? 」

위 문제의 적어도 한번 이상 $(6,6)$이 나올 가능성이 $\frac{1}{2}$보다 크려면 $\cdots$에서 제시된 $\frac{1}{2}$은 게임의 공정함을 의미한다. 공정한 게임을 진행하기 위해서 게임을 시작하기 전에 주사위를 던지는 횟수를 결정하는 이 문제에서, 문제를 해결하는 기준은 바로 게임의 공정함을 나타내는 수인 $\frac{1}{2}$에 있다. 즉, 주사위 게임의 공정함을 유지하기 위한 이 문제에서 주사위를 n번 던질 때 적어도 한 번 이상 $(6,6)$의 눈이 나올 확률 p를 여사건의 확률을 이용하여 구하면 $p = 1 - (\frac{35}{36})^n$이다. 따라서 $n = 24$일 때, $p = 0.4914$이고, $n = 25$일 때, $p = 0.5055$이므로, 적어도 한번 이상 $(6,6)$이 나올 가능성이 $\frac{1}{2}$보다 크게 하기위해서는 2개의 주사위를 최소한 25번 던져야 한다.

2.2 Pascal과 Fermat의 기댓값 계산

우연 게임에 대한 두 번째 문제는 이탈리아 수학자들 사이에서 다루어졌던 상금의 분배 문제이다.

이 분배 문제의 해법에서 Pascal은 한 게임을 이길 승률이 $\frac{1}{2}$ 로 같은 두 참가자 A, B가 게임이 계속된다고 할 때, n번의 게임을 이기기 위해 필요한 게임의 수를 이용하여 계산한 각 참가자의 승률에 비례해서 나누어 가져야 할 몫을 구하는 것을 내기 돈을 분배하는 원칙으로 정하였다.

Pascal은 분배문제를 처음에는 귀납적 방법으로 해결하였고, 이 후 Fermat와의 서신 교환이 이루어지는 가운데 수삼각형을 사용함으로써 일반적인 해법을 찾을 수 있었다. 이 문제에서 상금의 공정한 분배에 관한 수학적인 해석이 이루어지면서 기댓값의 개념은 발전하게 된다.

Pascal은 내기 돈을 분배하는 방법을 각 참가자가 이기기 위해 필요한 게임의 수로 설명하고 있다. 그는 A가 더 이겨야 할 게임 수가 a, B가 더 이겨야 할 게임 수는 b인 상황

에서 게임이 중단되었다고 할 때, 전체 내기 돈 중에서 A의 기댓값을 $e(a,b)$로 나타냈다. $e(a,b)$에 대한 Pascal의 계산은 <표Ⅱ-1>과 같다.

<표Ⅱ-1>에서 [1],[2]는 $e(0,n)=1, e(n,n)=\frac{1}{2}, n=1,2,\cdots$과 같이 나타낼 수 있다. $e(0,n)$는 A가 n번의 게임을 이기기 위해서 더 이겨야 할 게임 수가 0개이고, B가 더 이겨야 할 게임 수는 n개일 때 A의 기댓값이므로 1이다. $e(n,n)$는 A가 이겨야 할 게임 수는 n개이면서 마찬가지로 B가 이겨야 할 게임 수도 n개일 경우의 기댓값이다. 이 경우에 한 게임을 이길 승률이 $\frac{1}{2}$로 같은 두 참가자 A, B의 기댓값은 역시 같을 것이므로, 이 자연수일때 $e(n,n)=\frac{1}{2}$이다.

표 1: $e(a,b)$에 대한 Pascal의 계산

	A의 승패	a	b	$e(a-1,b)$	$e(a,b)$
				$e(a,b-1)$	
[1]		0	n		1
[2]		n	n		$\frac{1}{2}$
[3]		1	2		$\frac{1}{2}\times(1+\frac{1}{2})=\frac{3}{4}$
	A승	0	2	1	
	A패	1	1	$\frac{1}{2}$	
[4]		1	3		$\frac{1}{2}\times(1+\frac{3}{4})=\frac{7}{8}$
	A승	0	3	1	
	A패	1	2	$\frac{1}{2}+\frac{1}{4}=\frac{3}{4}$	
[5]		1	4		$\frac{1}{2}\times(1+\frac{7}{8})=\frac{15}{16}$
	A승	0	4	1	
	A패	1	3	[3]으로부터 $\frac{7}{8}$	
[6]		2	3		$\frac{1}{2}\times(\frac{1}{2}+\frac{7}{8})=\frac{11}{16}$
	A승	1	3	[3]으로부터 $\frac{7}{8}$	
	A패	2	2	[2]로부터 $\frac{1}{2}$	

[3]에서 [6]까지와 같이 Pascal은 $e(a,b)$의 계산에서 A가 이기거나 또는 A가 지는 두 가지 가능한 결과를 고려하는 조합적인 방법을 사용하였다. 우선 A가 이길 때의 $e(a-1,b)$와 B가 이길 때의 $e(a,b-1)$는 이전 단계를 이용하여 계산하였다. 그리고 어느 게임이든 A가 이기는 경우와 B가 이기는 경우인 2가지로 나눌 수 있고, 한 번의 게임에서 A나

B가 이기는 승률은 같으므로 $e(a,b)$는 $e(a-1,b)$와 $e(a,b-1)$의 평균, 즉 $e(a,b) = \frac{1}{2}[e(a-1,b)+e(a,b-1)]$의 방법으로 기댓값 $e(a,b)$을 계산하였다.

[3]에서 구한 $e(1,2)$을 예로 이 방법을 적용해 보면 $e(1,2)$는 A가 더 이겨야 할 게임의 수는 1개이고, B가 더 이겨야 할 게임의 수는 2개일 때 A의 기댓값이다. 이 상태에서 한 번 더 게임을 한다면 A가 이기는 경우와 B가 이기는 경우인 두 가지 경우만 나타난다.

이 때 $e(1,2)$에서 A가 이긴다면 A는 더 이겨야 하는 남은 게임은 없으며, B는 이기지 못하였으므로 $e(0,2)$이고, $e(0,2)=1$이다. 하지만 $e(1,2)$에서 B가 이긴다면 $e(1,1)$이고, $e(1,2)=\frac{1}{2}$이다. 따라서 $e(1,2)=\frac{1}{2}[e(0,2)+e(1,1)]=\frac{1}{2}\times(1+\frac{1}{2})=\frac{3}{4}$이다. 또한, Pascal은 a,b가 큰 값일 경우에도 $e(a,b)$의 계산이 가능한 일반적인 해법을 찾았는데, 그 방법은 기댓값 계산의 반복적인 과정을 수삼각형을 통한 계산과 대응시키는 방법이었다. Pascal이 제안한 대응식은 $e(a,b)=\frac{D_{a+b-1,b-1}}{D_{a+b-1}}$ 이다.

이 식에서 분자 $D_{a+b-1,b-1}$는 수삼각형에서 $a+b-1$번째 행의 숫자를 왼쪽부터 시작하여 b개의 숫자를 더한 값이고, 분모 D_{a+b-1}는 $a+b-1$번째 행의 숫자를 모두 더한 값으로 2^{a+b-1}와 같다. Pascal의 해법은 수학적 귀납법으로 다음과 같이 증명할 수 있다.

(1) $a+b=2$일 때 $e(0,2)=1, \quad e(2,0)=0, \quad e(1,1)=\frac{1}{2}$로 성립한다.

(2) $a+b=k$일 때, $e(a,k-a)=\dfrac{D_{k-1,k-a-1}}{D_{k-1}}$이라 하면,

$a+b=k+1$일 때,

$$\begin{aligned} e(a,k+1-a) &= \frac{e(a-1,k+1-a)+e(a,k-a)}{2} \\ &= \frac{1}{2}(\frac{D_{k-1,k-a}}{D_{k-1}}+\frac{D_{k-1,k-a-1}}{D_{k-1}}) \\ &= \frac{D_{k-1,k-a}+D_{k-a,a-1}}{2D_{k-1}} \\ &= \frac{D_{k,k-a}}{D_k} \end{aligned}$$

이므로, $a+b\geq 2$인 모든 자연수 n에 대해서, $a+b=n$일 때 Pascal이 제안한 대응식 $e(a,b)=\frac{D_{a+b-1,b-1}}{D_{a+b-1}}$은 성립한다.

또한 대응식 $e(a,b)=\frac{D_{a+b-1,b-1}}{D_{a+b-1}}$는 조합 기호와 Σ 를 사용하여 다음과 같이 나타낼 수 있다.

$$e(a,b)=\sum_{i=0}^{b-1}\binom{a+b-1}{i}\left(\frac{1}{2}\right)^{a+b-1}=\sum_{i=a}^{a+b-1}\binom{a+b-1}{i}\left(\frac{1}{2}\right)^{a+b-1}$$

<그림Ⅱ-1>에 나타난 Pascal의 수삼각형과 대응식을 이용하여 $a=2, b=3$인 경우에서 A의 기댓값 $e(2,3)$을 구하면, $e(2,3)=\frac{1+4+6}{2^{2+3-1}}=\frac{11}{16}$이다.

Pascal은 기댓값을 계산할 때, A가 이기거나, 또는 B가 이기고 A가 지는 두 가지 경우로 나누어 계산하는 조합의 방법을 사용하여 분배문제를 해결하였다.

반면에 Fermat의 해법은 각 참가자가 더 이겨야 할 게임수가 a, b만큼 남았을 경우에 게임을 많이 해야 $a+b-1$번만 하면 승패가 결정이 날 것이라는 것과 $a+b-1$번 게임에서 두 참가자가 이기고 지는 총 경우 중에서 A가 이기는 경우의 비율이 A가 가질 내기 돈의 비율인 기댓값이라는 사고를 기반으로 하고 있다.

1 1	→	1번째
1 2 1	→	2번째
1 3 3 1	→	3번째
1 4 6 4 1	→	4번째
1 5 10 10 5 1	→	5번째
···		···

그림 1: Pascal의 수삼각형

Fermat의 해법에 따라 $(a,b)=(2,3)$인 경우를 구해 보면, 2번 게임을 할 때의 경우의 수는 2^2가지이다. 즉 (승 승, 패 패), (승 패, 패 승), (패 승, 승 패), (패 패, 승 승) 중에서 A가 최종적으로 이기는 경우는 2번의 게임을 모두 이기는 1 가지 경우의 수가 있다. 그리고 같은 방법으로 3번 게임을 하면 2^3가지의 가능한 경우의 수 중에서 A가 이기는 경우가 2가지 있으며, 4번 게임을 하면 2^4가지의 가능한 경우의 수 중에서 A가 이기는 경우는 3 가지가 있다.

따라서 A의 기댓값 $e(2,3)=\frac{1}{2^2}+\frac{2}{2^3}+\frac{3}{2^4}=\frac{11}{16}$이 되는데, 이러한 Fermat의 해법을 현대적인 기호로 나타내면 $e(a,b)=\sum_{i=0}^{b-1}\binom{a-1+i}{a-1}\left(\frac{1}{2}\right)^{a+i}$이다. 즉, A의 기댓값 $e(a,b)$에 대한 Fermat의 해법은 게임의 참여자인 A나 B가 게임을 이기는 경우를 A 또는 B로 표현할 때, A를 $a-1$개, B를 i(단, $0\leq n\leq b-1$)개 배열한 다음에 마지막 자리에 A를 배열하는 것이다.

2.3 Huygens의 기댓값

기댓값(expectation)이란 용어를 최초로 사용한 호이겐스(Christian Huygens:1629 ~1695)는 1657년에 그 당시 확률 이론의 지침서가 되는「우연게임에 관한 추론(De Ratiociniis in Ludo Aleae)」을 출판하였는데, 이 논문에서 Huygens가 주장한 기댓값의 의미를 유추해 볼 수 있다.

그는 어떤 게임에 참가한 참가자들은 누구든지 게임에서 돈을 잃고 싶지는 않을 것이므로 공정한 게임에서는 모두가 같은 기회를 얻을 수 있다는 개념인 게임의 공정성을 염두에 두고, 기댓값 개념을 발전시켜 나갔다. Huygens는 기댓값에 관한 3개의 정리를 주장하였는데, Huygens가 언급한 3개의 정리는 다음과 같다.

정리 2.1. 어떤 게임에서 a를 얻을 수 있는 가능성과 b를 얻을 수 있는 가능성이 같다면, 이 게임은 $\frac{a+b}{2}$ 만큼의 가치가 있다.

정리 2.1의 증명: 두 명의 참가자가 참가하여 지면 a를 얻고, 이기면 b를 얻는 복권추첨에서, 각 참여자가 낸 판돈이 x라고 하고, a또는 b를 얻을 가능성은 같다고 하자. 승자는 전체 판돈 $2x$을 따서 그 중에서 a만큼을 패자에게 주기로 한다면 승자는 $2x-a$만큼을 갖게 되는 것이다. 따라서 $2x-a=b$이고, $x=\frac{a+b}{b}$이다.

정리 2.2. 어떤 게임에서 a또는 b또는 c를 얻을 수 있는 가능성이 각각 같다면, 이 게임은 $\frac{a+b+c}{3}$ 만큼의 가치가 있다.

정리2의 증명: 세 명의 참가자가 참가하여 지면 a또는 b를 얻고, 이기면 c를 얻는 복권추첨에서, 각 참가자가 낸 판돈이 x라고 하고, a또는 b 또는 c를 얻을 가능성은 같다고 하자. 승자는 전체 판돈 $3x$을 따서 그 중에서 a또는 b만큼을 패자에게 주기로 한다면 승자는 $3x-c$만큼을 갖게 되는 것이다. 따라서 $3x-c=a+b$이고,$x=\frac{a+b+c}{3}$이다.

정리 2.3. 어떤 게임에서 a를 얻을 수 있는 가능성이 p, b 를 얻을 수 있는 가능성이 q라면, 이 게임은 $\frac{aq+bp}{p+q}$ 만큼의 가치가 있다.

정리 3의 증명: $p+q$명의 참가자가 참가하여 지면 a를 얻고, 이기면 b를 얻는 복권추첨에서 각 참가자가 판돈 x를 지불하고, a또는 b를 얻을 가능성은 같다고 하자. 전체 참가자 중에서 p명이 이기고, q명은 지게 된다면 이긴 참가자들은 전체 판돈 $(p+q)x$ 중에서 aq만큼을 뺀 나머지를 p명이 나누어 갖게 될 것이다. 따라서 $b=\frac{(p+q)x-qa}{p}$이므로, $x=\frac{aq+bp}{p+q}$이다.

정리에 나타난 Huygens의 기댓값에 대한 기본적인 사고는 각 참가자가 게임에서 이기거나 질 가능성이 동등하다고 가정하면서 같은 양의 돈 n만큼을 내고 참여하는 우연 게임의 한 형태인 복권의 추첨이었다.

동등한 가능성을 갖는 복권 추첨에서 게임의 가치로 표현되는 참가자의 기댓값은 참가자가 게임에 참여하기 위해 내야하는 참가자마다의 판돈으로 정의된다. 이를 통해 동등한 기회를 갖는 유한 명의 참가자가 복권 추첨을 한다면, 이 게임에서의 기댓값은 상금의 산술 평균과 같게 되고, Huygens가 주장한 게임의 가치에 대한 의미는 지금 기댓값의 정의와 같다는 것을 알 수 있다.

2.4 페테르부르크의 역설과 Montmort의 기댓값

성 페테르부르크 학술원에서 니콜라스 베르누이(Nicholas Bernoulli:1687～1759)가 제안한 기댓값에 대한 한 문제가 이후에 ∞의 기댓값이 존재하는 시행이 있을 수 있다는 페테르부르크 역설(Petersburg paradox)로 알려지게 되었다.

다음은 Nicholas Bernoulli가 제시한 페테르부르크 역설의 초기 문제이다.

「페테르부르크 역설의 초기 문제 : 주사위를 던져 처음에 6의 눈이 나오면 A가 B에게 동전 1개를 주고, 두 번째에 6의 눈이 나오면 A가 B에게 동전 2개를 주고, 세 번째에 6의 눈이 나오면 A가 B에게 동전 3개를 주고, 네 번째에 6의 눈이 나오면 A가 B에게 동전 4개를 주는 규칙으로 게임을 무한히 계속한다면, B의 기댓값은 얼마인가?」

Nicholas Bernoulli는 활동했던 당시에는 우연 게임에서 그 게임이 공평하게 이루어지기 위한 기댓값을 조합적인 사고를 통해 계산하는 방법이 많이 쓰였는데, 다음 풀이는 Nicholas Bernoulli가 조합적인 방법으로 계산했던 B의 기댓값을 구하는 풀이와 급수를 이용한 방법으로 B의 기댓값을 구한 것이다.

Nicholas Bernoulli의 풀이 : B의 기댓값을 x, 첫 번째 시행에서 6의 눈이 나오지 않았을 때 B의 기댓값을 y라 하면 $x = \frac{1}{6} + \frac{5}{6}y$.

또한 주사위를 무한히 계속해서 던지는 게임이므로 처음에 주사위를 던져 6의 눈이 나오지 않은 후에 B는 $2, 3, 4, 5, 6$의 순서로 동전을 받는 것을 기대할 수 있고, 원래 동전을

지급하는 순서는 $1, 2, 3, 4, 5$로 각 수열의 항은 1만큼의 차이가 있으므로

$$y = x + 1 \cdots\cdots ①$$

①을 $x = \frac{1}{6} + \frac{5}{6}y$에 대입하면, $x = \frac{1}{6} + \frac{5}{6}(x+1)$. 계산하면 $x = 6$이다.

급수를 이용한 풀이: n번째 시행까지의 기댓값을 S_n이라 하면

$$S_n = \frac{1}{6} - 1 + \frac{5}{6} - \frac{1}{6} - 2 + \frac{5}{6} - \frac{5}{6} - \frac{1}{6} - 3 + \cdots \left(\frac{5}{6}\right)^{n-1} - \frac{1}{6} - n \cdots\cdots ①$$

①$\times \frac{5}{6}$에서

$$\frac{5}{6}S_n = \frac{5}{6} - \frac{1}{6} - 1 + \frac{5}{6} - \frac{5}{6} - \frac{1}{6} - 2 + \frac{5}{6} - \frac{5}{6} - \frac{5}{6} - \frac{1}{6} - 3 + \cdots \left(\frac{5}{6}\right)^{n} - \frac{1}{6} - n \cdots\cdots ②$$

① − ②에서

$$\frac{1}{6}S_n = \frac{1}{6} + \frac{5}{6} - \frac{1}{6} + \left(\frac{5}{6}\right)^2 - \frac{1}{6} - 3 + \cdots + \left(\frac{5}{6}\right)^{n-1} - \frac{1}{6} - \left(\frac{5}{6}\right)^{n} - \frac{1}{6} - n \cdots\cdots ③$$

③에서 무한급수의 합을 구하면 $\lim_{n\to\infty} S_n = 6 - \frac{5/6}{1-(1/6)} = 6$ 이므로 B의 기댓값은 6이다.

이후 발전된 형태의 페테르부르크 역설의 문제를 정리하면 다음과 같다.

> 「페테르부르크 역설: 앞면과 뒷면이 나오는 정도가 같은 공정한 동전을 사용하여 앞면이 나올 때까지 무한히 계속해서 던지는 게임에서 만약 x번째에서 앞면이 나온다면 B가 A에게 2^x만큼의 금액을 받는다고 할 때, B의 기댓값은 얼마인가?」

발전된 형태의 페테르부르크 역설에서 제시한 기댓값을 무한급수의 합을 이용하는 방법으로 계산해 보면 아래와 같다.

풀이: n번째 시행까지의 B의 기댓값을 S_n이라 하면

$$S_n = \frac{1}{2} - 2^1 + \left(\frac{1}{2}\right)\frac{1}{2} - 2^2 + \left(\frac{1}{2}\right)^2\frac{1}{2} - 2^3 + \cdots + \left(\frac{1}{2}\right)^{n-1} - \frac{1}{2} - 2^n$$

$$= 1 + 1 + 1 + \cdots + 1 = n$$

이므로 무한급수의 합을 구하면 $\lim_{n \to \infty} S_n = \infty$ 이다.

따라서 B의 기댓값은 무한이다.

B의 기댓값은 수학적으로 정확한 계산이 가능함에도 불구하고, 수학적으로 계산된 B의 기댓값 다시 말해 A가 지불해야 할 무한의 금액과 공정한 게임이 되도록 B가 게임에 참가하기 위해서 지불해야 하는 무한의 금액에 대해서 무한 개념이 명확히 정립되지 못했던 그 당시의 수학자들이 심정적으로 수용할 수 없었다는 것에서 페테르부르크 역설이 생기게 되었다.

즉 논리적으로 정확한 수학적 계산에 의해 나오는 기댓값의 결과를 실제로는 현실에 적용할 수 없다는 모순된 결과에서 제기되는 페테르부르크 역설은 실생활에서는 무한 번의 게임을 개최할 수 없고, 따라서 무한 횟수에 대한 기댓값을 계산할 수 없을뿐더러 이 기댓값을 게임의 공정한 참가비용으로 제시하는 것은 타당하지 못하다는 것이다. 따라서 페테르부르크 역설은 수학적 기댓값이 실생활에서 합리적인 예측 수단이 항상 될 수는 없다는 것을 보여주는 예이다.

오늘날 우리가 사용하는 기댓값의 의미는 몽모르(Pierre Raymond de Montmort :1678 ~1719)에 의해 완성되었다고 할 수 있다. Montmort가 저술한「우연게임의 분석에 관한 에세이(Essay d' analyse sur les jeux de hazard)」에서, 그는 어떤 게임에서 이길 확률을 p, 전체 판돈을 S라고 했을 때, 참가자의 기댓값을 $e = pS$라고 정의하였고, S를 얻을 기회가 m번이고, 아무 것도 얻지 못할 기회가 n번일 때, 참가자의 기댓값 $e = \frac{mS + n \times 0}{m+n}$ 이라고 주장하였다. 또한 어떤 게임에서 참가자들이 게임에 참가하기 위해서 제공한 참가금의 비가 그 참가자들의 기댓값의 비와 같을 때, 그 게임은 공정하다고 하면서 기댓값에서 판돈을 뺀 만큼을 참가자들이 가지게 될 이익으로 보았다.

3 결론

형식화되기 이전의 실험적이며 귀납적인 수학, 통찰에 의해서 제시된 발생 상태 그대로의 수학, 발명되고 논쟁을 통해 변화되고 발전되면서 정선되어 가는 수학속의 역사에 대한 연구는 수학적 지식이나 주요한 아이디어가 그 시기에 발생하게 된 이유를 찾고, 그 개념이 전개되고 발전해 나가면서 직면했던 여러 가지 도전들을 직접 체험해보면서, 수학자들이 처음부터 오랜 시기를 거치면서 이루었던 수학적 사고의 과정을 다시 경험해 본다는 데 의

의가 있다.

기댓값은 확률론의 핵심적인 내용 중 하나이지만 비교적 단순하게 취급되어 온 것이 사실이다. 특히 학교수학에서 일반적으로 기댓값 개념을 다룰 때, 몇 가지 예를 통해 기댓값을 단편적으로 이해한 후 바로 응용문제의 풀이를 시작함으로 인해 기댓값의 정확한 의미를 이해하기 어렵다.

이에 본 연구는 확률 개념의 기원인 여러 가지 우연 게임과 함께 기댓값과 관련이 있는 공정한 분배의 의미가 발달되어온 역사에 대해 간략하게 살펴보았다. 그리고 확률론 연구의 직접적인 동기가 되었던 Fermat와 Pascal의 서신에 나타난 공정한 분배의 문제, 반세기 동안 확률론의 지침서가 되어왔던 Huygens의 저서에 나타난 기댓값과 관련된 정리와 증명, Nicholas Bernoulli의 페테르부르크 역설, 그리고 Montmort의 현대적인 기댓값 정의에 대해서 살펴보면서 기댓값의 개념 및 그 변화에 기여한 역사적 사건들이 갖는 의의와 기댓값의 의미를 재조명하고자 하였다.

참고 문헌

[1] 고민지, 《확률 역사 속에 나타난 Huygens의 14 Proposition》, 수학사랑, 제8회 Math Festival, 293–304,2006.

[2] 노영석, 《PASCAL과 확률이야기》, 통계학술논집, 10(1995), 1–49.

[3] 박성열, 《기대값에 관하여》, 단국대학교 수학교육 석사학위 논문1998.

[4] 박영희, 《통계 영역에서 대푯값의 의미와 지도에 관한 고찰》, 학교수학, 3(2001), No.2, 281–293.

[5] 백경호, 《고등학교 확률과 통계 영역에서 현실적 수학교육의 적용》, 한국교원대학교 석사학위 논문, 2004.

[6] 신보미, 《시뮬레이션을 활용한 확률 지식의 교수학적 변환 방식》, 한국교원대학교 박사학위 논문, 2007.

[7] 이금아, 《Excel을 활용한 조합문제 해결과 확률분포 시뮬레이션》, 청람수학교육, 15(2004), 217–240.

[8] 이경화, 《확률개념의 교수학적 변환에 관한 연구》, 서울대학교 박사학위논문, 1996.

[9] 이석훈, 김응환, 《통계와 확률 지도론》, 서울 : 경문사,2003.

[10] 장대흥, 이효정, 《제7차 수학과 교육과정에 따른 1～10단계 확률 및 통계단원 분석》, 응용통계연구, 18(2005), No. 1, 229-249.

[11] 장인홍, 《고전확률론과 중심극한정리에 대한 역사적 고찰》, 한국수학사학회지, 15(2002), No.3, 65–74.

[12] 조규생, 《「확률과 통계」 교과서 단원 전개에 관한 연구》, 영남대학교 박사학위 논문,2002.

[13] 조재근, 《생물측정학–멘델주의 논쟁에 대한 통계학사적 고찰》, Communication of the korean

Statistical Society, 15(2008), No.3, 303–324.

[14] 허민, 《수학교육에 활용할 옛 문제 연구》, 한국수학사학회지, 13(2008), No. 1, 33–48.

[15] Stephen, M. Stigler, *The history of statistics*, 1986, 조재근(역), 『통계학의 역사』, 서울 : 한길사, 2005.

[16] Anders Hald, *History of Probability and Statistics and Their Applications before 1750*, New Jersey: wiley, 2003.

[17] Arthur, B., *The early history of average values and implications for education*, Journal of Statistics Education, Vol, 11(2003), No.1.

[18] Brousseau, G., *Theory of didactical situations in mathematics*, Edited and traslated by N. Balacheff, M. Cooper, R. Sutherland & V. Warfield, Dordrecht: Kluwer academic publishers, 1997.

[19] David Salsburg, *The lady tasting tea*, 2001. 최정규(역), 『천재들의 주사위 — 20세기를 만든 통계학의 혁명들』, 서울 : 뿌리와 이파리, 2003.

[20] Gábor J. Székely, *Paradoxes in Probability Theory and Mathematical Statistics*, D. Reidel Publishing Company, 1986.

[21] I. B. Cohen, *The triumph of number*, 2005. 김명남(역), 『세계를 삼킨 숫자이야기』, 서울 : 생각의 나무, 2007.

[22] Richard J. P., *Correspondence of Nicolas Bernoulli concerning the St Petersburg Game*, 1999.

Historical reflections on the expectation

Daejeon Songchon High School Lee Jong-Hak

In this paper we study the expectation from the past which would be based in the initial Probability Theory. These study can show the value of expectation in the initial concept of Probability and illuminate the concept being taught.

Key Words: Probability, expectation.

2000 Mathematics Subject Classification : 97-03
ZDM Subject Classification : A34

접수일 : 2010년 6월 17일 수정일 : 2010년 7월 30일 게재확정일 : 2010년 8월 3일