韓國數學敎育學會誌 시리즈 E
<數學敎育 論文集>
제 9집, 1999. 7. 299-316.

J. Korea Soc. Math. Ed. Ser. E:
Communications of Mathematical Education
Vol. 9, July. 1999. 299-316.

확률 통계 지도를 위한 Excel의 활용방안

서 현 경 (광주양산중학교)
강 순 자 (전 남 대 학 교)
임 해 경 (광주교육대학교)

본 논문에서는 스프레드쉬트 프로그램 중에서 가장 활용도가 높은 Excel을 이용하여 만든 여러 가지 모의실험이 확률 통계학습에 어떻게 활용되는 지를 제시함으로써 개념의 지도 및 문제풀이 능력 향상의 효율성을 높이는 방안을 모색하고자 한다. 즉 이는 단순한 이론적 수치계산이 아닌 구체적 경험을 제시하여 학생들에게 확률적 상황에 내재된 확률적 정보의 의미를 파악하게 함으로써 확률의 개념에 대한 이해를 돕고 확률 통계단원에 대한 흥미를 유발케 하고자 한다.

1. 서론

최근에 우리는 학교정보화라는 이름으로 인터넷을 통한 교육이나 원격교육에 관한 많은 사례들을 접하고 있다. 또한 과학정보통신의 시대에 부합하여 PC의 보급이 보편화되면서 교육현장에서도 컴퓨터를 이용한 효과적인 지도 방안들이 연구되어지고 있다.

컴퓨터를 이용한 수학교육 관련 연구들은 주로 기호계산 및 그래픽 기능이 뛰어난 Mathematica, Maple, Geometer' s sketchpad, CabriII 등을 함수, 방정식, 기하학에 활용하는 방안들이 대부분이지만 수학에서도 실험학습이 가능하다는 사실 또한 보여주고 있다. 특히 최근에 들어서는 쉽게 접할 수 있고 비교적 활용하기가 쉬운 스프레드쉬트(spreadsheet) 관련 프로그램(Excel, Lotus-123, 하나쉬트, 마이쉬트등)을 이용한 수학교육에 관한 연구가 이루어지고 있다(장경윤, 1997; Deane, 1984)

확률과 통계는 중학교 2학년 교과서에 처음 소개되는 개념이다. 물론 저학년에서부터 이와 관련된 여러 가지 기초개념 들을 학습하여 왔지만 독립적인 단원으로 소개되는 것은 중학교 수준에서이다. 확률 개념은 정보화 사회에서 많은 정보를 분석하고 불확실한 현상을 모형화하여 미래를 예측하는 통계의 밑바탕을 이루고 있어 현대사회에서 매우 중요한 역할을 하고 있고 전 세계적으로 수학교육과정에서 확률의 중요성이 강조되어 왔다. 그럼에도 불구하고 확률 개념 자체가 애매하고 이해하기 쉽지 않으므로 확률 및 확률분포 개념의 학습 지도가 결코 용이하지 않다. 특히 주변의 여러 현상들이 일어날 가능성에 대해 추론하여 예측하는 것인 만큼 다른 단원에 비해 실생활과 더욱 관련이 깊은 단원인데도 구체적인 실험이나 개념에 대한 다각적인 접근 없이 수치적으로 계산하는 것만 주로 다루어져 옴으로써 확률에 대한 학생들의 오개념의 근원이 되고 있으며 흥미를 잃게하는 원인이 된

다는 비판을 받아왔다.

확률 통계 단원은 다른 영역과는 다른 귀납적인 특성을 가지고 있기 때문에 논리적 추론 보다는 실제 실험적 상황에서 구체적 경험을 통하여 학생들이 그 의미를 이해하도록 하는 게 바람직하다. 특히 고등학교에서 학생들은 확률분포에 대한 개념과 활용에 대해 학습해야 하는데 이 학습은 이항분포등을 포함하는 특수한 분포 및 정상분포에 대한 도입을 포함한다. 학생들은 사건의 발생가능성을 예견하거나 어떤 표본 집단에 근거한 예측의 질을 평가하기 위하여 이들 분포를 이용해야 한다 (NCTM, Draft, 1998). 결국 고등학교 확률 통계수업에서 학생들이 확률의 정의를 바탕으로 한 확률변수와 확률분포의 개념을 얼마만큼 이해하고 있는가 하는 것이 중요한 문제가 된다. 그것은 확률분포에 대한 올바른 개념 정립이 통계학적 추론의 개념적 기초를 형성하기 때문이다.

학생들은 자료 분석, 통계 그리고 확률을 공부하는 기간 내내 개념을 제시하거나 계산을 수행하거나 자료에 대한 표현을 창조하거나 모의 실험으로부터 데이터를 제공하는 등의 학습 활동을 위하여 휴대용 계산기나 컴퓨터를 이용할 수 있으며(NCTM. Draft, 1998), 따라서 이제 어떻게 컴퓨터를 효과적으로 이용할 것인가하는 방법론적인 문제가 제기된다.

스프레드쉬트는 도구 소프트웨어의 하나로써 행과 열로 짜여진 작업지 위에 수량적 자료처리(사칙연산, 정렬)와 그래픽을 가능하게 한 통합 프로그램으로 논리 및 수학함수를 내장하고 있어서 다양한 연산기능 뿐 아니라 여러 가지 모의 실험 등의 활동을 용이하게 할 수 있다는 장점이 있다. 따라서 이러한 스프레드쉬트의 장점을 이용하여 여러 가지 모의 실험을 경험해 본다면 학생들이 주어진 현상을 이해하는데 큰 도움이 될 수 있을 것이다.

따라서 본 고에서는 스프레드쉬트 프로그램 중에서 가장 활용도가 높은 Excel(스프레드쉬트 프로그램 사용자중 59.1%가 사용하고 있음; http://arom.etri.re.kr/htdocs/data/ko-2-1.html)을 이용하여 확률 통계학습에 적절히 활용할 수 있는 여러 가지 모의실험을 제시함으로써 확률 및 확률 분포의 개념 지도의 효율성을 높이고자 한다. 즉 이는 단순한 이론적 수치계산이 아닌 구체적 경험을 제시하여 학생들에게 확률적 상황에 내재된 확률적 정보의 의미를 파악하게 함으로써 확률의 개념에 대한 이해를 돕고 확률 통계단원에 대한 흥미를 유발케 하고자 한다.

2. 컴퓨터 모의 실험

많은 확률 통계교육에 관한 연구보고서에서는 확률과 통계의 교육에 모의실험의 도입은 학생들의 과목에 대한 흥미를 유발시키고 탐구욕을 촉진시키는 등 학습효과를 가져온다고 말하고 있다 (Dambolena, 1986).

확률 통계교육에서의 모의 실험은 상황에 대한 보다 깊은 이해를 갖게 해 주고 그것에 포함된 수학적인 것에 대해서도 이해하게 해 준다. 컴퓨터 모의 실험은 여러 가지 교육 목적을 촉진시키기 위해 사용될 수 있는 데 과목에 대한 학생들의 흥미를 유발하고 탐구욕을 촉진 시키고 개념을 깊게 이

해시키고 더불어 모의실험 기법을 숙달케 한다(Coburn et al,1985).

예를 들어 우리 주변에는 많은 불확실한 사건들이 반복적으로 일어나는 경우가 있다. 동전을 반복적으로 던진다든지 주사위를 계속 같은 방법으로 굴린다든지 하는 것이다. 동전의 앞면이 나올 확률은 수학적 확률로는 1/2이다. 그러나 상대 빈도 개념(relative frequency concept)에서는 이 실험을 수없이 반복하였을 때 전체 실험횟수에서 앞면이 나오는 빈도수가 1/2 에 접근한다는 의미이다. 즉 n번의 시행횟수에서 어떤 사건이 일어날 횟수를 r이라 하면 상대빈도 r/n은 시행횟수 n을 무한대로 함에 따라 어떤 극한값에 가까워진다는 것이다. 이러한 방법으로 확률을 측정하기 위해서는 많은 경우를 관찰하여야 한다.

모의 실험을 통한 관찰은 확률적 상황과 수학적 모델간의 관련성을 인식하고 확률의 개념을 음미하도록 자극하는 데 필요하다. 모의 실험은 원판 돌리기, 동전 던지기, 주사위 던지기, 난수표 사용 등과 같은 물리적 모의 실험과 오늘날 수업에서 쉽게 이용될 수 있는 컴퓨터 모의 실험이 있다. 물리적 모의 실험은 시간을 많이 필요로 하는 반면 컴퓨터 모의 실험은 짧은 시간 내에 할 수 있으면서도 많은 장점을 지니고 있다.

첫째, 반복횟수를 늘릴 수 있으므로 결과의 불확실성과 변화를 감소 시킬 수 있으며, 새로운 종류의 패턴을 발견할 수 있다.

둘째, 모델의 사정을 바꾸고, 실험을 세련시키고, 구해진 자료를 다른 방법으로 분석함으로써 연구영역을 확장할 수 있다.

셋째, 모델, 통계적 처리과정, 그래프 등을 보다 효과적으로 제공함으로써 유연한 표현이 가능하다(Biehler, 1991).

컴퓨터 모의 실험은 확률적 실험과 이론 사이의 차이를 좁혀주고 현실적 문제 속에서 확률적 이론을 발견하는 데 도움을 준다.

3. 확률·통계 지도를 위한 Excel의 활용방안

이제 확률·통계 지도를 위해 Excel의 여러 기능들을 이용한 활용 방안들을 예시하려고 한다.

활용1. 수학적 확률과 통계적 확률의 관계

일반적인 독립시행에서 성립하는 큰 수의 법칙은 어떤 시행에서 사건 A가 일어날 확률을 p라 하고 이 시행을 n회 반복했을 때 A가 일어나는 횟수를 x라고 할 때, n을 크게 하면 상대도수 x/n는 확률 p에 한없이 가까워진다는 것이 거의 확실하다는 것이다.

즉, 한 개의 동전을 던질 때 나올 수 있는 모든 경우는 앞면 또는 뒷면의 2가지이다. 또한 한 개의 주사위를 던질 때 나올 수 있는 모든 경우는 1, 2, 3, 4, 5, 6 의 6가지 경우이다. 앞에서 살펴보았듯이 중·고등학교 교과과정에서의 확률은 주로 수학적 확률에 관한 이야기이다. 더욱이 실제 실험 상

황에서 얻어지는 통계적 확률에 관한 이야기는 중학교 과정에서 거의 언급하고 있지 않으며 통계적 확률과 수학적 확률사이의 연결 관계에 대해서는 고등학교 과정에서 큰 수의 법칙으로 잠깐 언급이 되어 있을 뿐 학생들이 직접 수업시간에 동전이나 주사위를 던져보아 그 결과를 관찰함으로써 확률 개념의 여러 정의들과 그들 사이의 관계를 살펴보도록 할애하고 있지 않다. 사실 현실적으로 동전이나 주사위로 모의 실험을 해 볼만큼 수업시간이 충분하지 않은 것이 사실이다.

이런 상황에서 학생들은 수학적 확률을 정의하기 위한 배경 조건에 대해서 충분히 생각해 보지 않은 상태에서 무작정 확률을 계산하는 방법만을 배우게 된다. 따라서 학생들 중에는 수치로 나타내지는 수학적 확률에 대해서 더러 잘못 이해하고 있는 것도 발견할 수 있다(특히 확률의 기초개념을 배우는 중학생인 경우에 많다). 예를 들면 어떤 사건이 반드시 주어진 확률에 비례하여 일어난다고 생각하는 경우가 있다. 즉 동전 1개를 던질 때 앞면이 나오는 확률은 1/2이므로 동전을 두 번 던지면 그 중에 한 번은 반드시 앞면이 나온다고 생각하는 것이다. 실제로 동전 1개를 두 번 던져보면 두 번다 앞면이 나올 수도 있고 두 번 다 뒷면이 나올 수도 있는 가능성을 잊고 있다.

또한 수학적 확률을 정의하기 위해서는 한 개의 동전을 던질 때 앞면과 뒷면이 동시에 나타날 수 없듯이 '서로 반대되는 사건이 동시에 일어나지 않는다'라는 조건과 앞면과 뒷면이 일어날 가능성이 같아야 하듯이 '일어날 가능성이 같다'라는 조건이 필요하다. 그러나 실제 상황에서는 이러한 조건을 만족시키지 않는 경우가 더 많다. 즉 동전 1개를 10번 던져서 나온 결과를 보면 10번 중 앞면이 7번 뒷면이 3번 나오는 것처럼 앞면과 뒷면이 나오는 가능성이 각각 다른 경우가 더 많기 때문이다.

확률의 정의 자체에서 오는 위와 같은 여러 가지 논쟁점들은 아직도 수학자들 사이에서 확률의 정의에 대한 의견의 대립으로 남아 있다. 여기서 학생들로 하여금 컴퓨터를 이용하여 동전이나 주사위 던지기 등과 같은 실험을 하여 그 결과를 분석해 보고 수학적 확률과 통계적 확률사이의 관계에 대해서 알아볼 수 있는 경험을 제공한다면 확률 개념의 정의에 대한 보다 깊은 이해를 도울 수 있을 것이다. Excel은 프로그램이 쉽고 복사기능이 있어 같은 실험을 여러 번 반복하여 시행하는 경우 몇 초 동안에 충분한 횟수만큼 동전을 던져 보는 실험을 해 볼 수 있어서 수업시간에 충분히 활용할 수 있다.

다음 예들은 Excel로 작성한 프로그램을 가지고 여러 가지 모의실험을 하여 관련된 문제를 해결하면서 확률개념에 관하여 자연스럽게 학습할 수 있도록 제시되었다.

[문제] 한 개의 동전을 던질 때, 앞면과 뒷면이 나올 가능성이 서로 같은가? 실제로 여러 번 던져서 앞면과 뒷면이 나타나는 횟수를 조사하여 보아라.

이것은 확률의 뜻에 대해서 설명하기 전에 중학교 과정에서 처음 도입 부분에 제시되는 질문이다. '실제로 여러 번' 이라는 말이 나왔지만 앞면과 뒷면이 나올 가능성을 비교해 보기 위해서는 충분히 많은 횟수(적어도 30회 이상)만큼 실험을 해 보아야 할 것이다.

다음은 위 문제에 대한 답을 하기 위하여 Excel을 이용해서 모의실험을 한 것이다.

	A	B	C
1	동전1개 던지기		
2	시행횟수	난수생성	나온 결과
3	1	=RANDBETWEEN(0,1)	=IF(B3=0,"앞","뒤")
4	2	=RANDBETWEEN(0,1)	=IF(B4=0,"앞","뒤")
5	3	=RANDBETWEEN(0,1)	=IF(B5=0,"앞","뒤")
6	4	=RANDBETWEEN(0,1)	=IF(B6=0,"앞","뒤")
7	5	=RANDBETWEEN(0,1)	=IF(B7=0,"앞","뒤")

<그림 1>

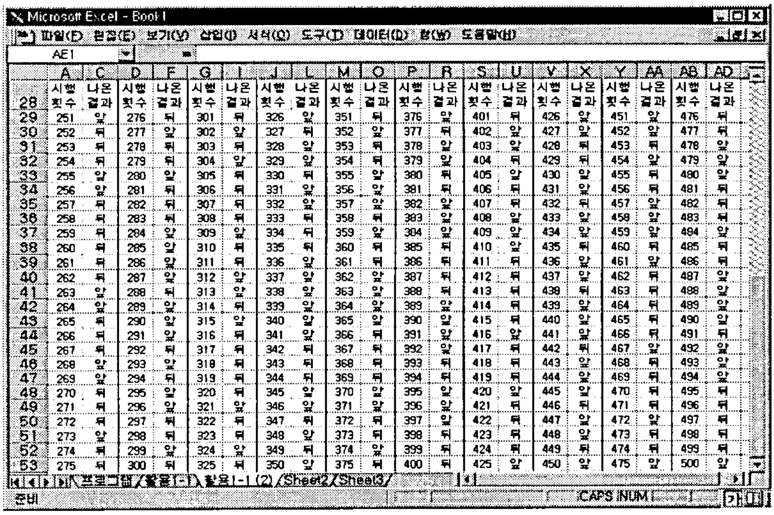
이것은 모의실험을 위한 프로그램인데 여기서 쓰여진 Excel language 는

· = RANDBETWEEN(0,1) : 동전을 던지면 나올 수 있는 경우는 앞면과 뒷면의 2가지이므로 0과 1중 두 숫자 가운데 임의로 난수를 발생시킨 것이다. 문자의 형태로 "앞"과 "뒤" 중 임의로 발생시키면 좋겠지만 이런 방법은 없으므로 "앞"을 대변하는 숫자 0과 "뒤"를 대변하는 숫자 1을 가지고 난수를 발생시켰다.

· = IF(B3=0, "앞", "뒤") : B열3행의 수가 0이면 "앞"을 1이면"뒤"를 인쇄하도록 한 것이다. 여기서 B3 은 상대주소로써 나중에 4행 이하의 프로그램은 따로 쓰지 않고 3행의 내용을 블록 설정한 후 마우스를 드래그 함으로써 복사하면 B4, B5, B6, … 등으로 행마다 변환하여 복사가 된다(500회 던진 모의실험인 경우 499개의 행에 복사한다).

	A	C	D	F	G	I	J	L	M	O	P	R	S	U	V	X	Y	AA	AB	AD
1	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과
2	1	뒤	26	앞	51	앞	76	뒤	101	앞	126	앞	151	앞	176	앞	201	뒤	226	뒤
3	2	앞	27	앞	52	앞	77	뒤	102	뒤	127	뒤	152	뒤	177	뒤	202	앞	227	앞
4	3	앞	28	뒤	53	앞	78	뒤	103	앞	128	뒤	153	뒤	178	앞	203	뒤	228	앞
5	4	뒤	29	앞	54	뒤	79	앞	104	앞	129	뒤	154	앞	179	뒤	204	앞	229	뒤
6	5	뒤	30	앞	55	뒤	80	뒤	105	뒤	130	앞	155	뒤	180	뒤	205	뒤	230	앞
7	6	뒤	31	뒤	56	뒤	81	뒤	106	앞	131	뒤	156	뒤	181	앞	206	뒤	231	앞
8	7	앞	32	앞	57	뒤	82	앞	107	뒤	132	뒤	157	뒤	182	앞	207	앞	232	앞
9	8	앞	33	앞	58	앞	83	뒤	108	앞	133	앞	158	앞	183	앞	208	뒤	233	뒤
10	9	앞	34	뒤	59	앞	84	앞	109	앞	134	앞	159	앞	184	앞	209	앞	234	뒤
11	10	뒤	35	뒤	60	뒤	85	뒤	110	앞	135	뒤	160	뒤	185	뒤	210	앞	235	뒤
12	11	뒤	36	뒤	61	앞	86	앞	111	뒤	136	앞	161	앞	186	앞	211	앞	236	앞
13	12	뒤	37	뒤	62	뒤	87	앞	112	뒤	137	뒤	162	앞	187	앞	212	앞	237	앞
14	13	앞	38	앞	63	앞	88	뒤	113	앞	138	뒤	163	앞	188	앞	213	뒤	238	앞
15	14	뒤	39	앞	64	앞	89	앞	114	앞	139	앞	164	뒤	189	앞	214	앞	239	앞
16	15	뒤	40	뒤	65	뒤	90	뒤	115	뒤	140	앞	165	뒤	190	뒤	215	뒤	240	앞
17	16	뒤	41	뒤	66	뒤	91	앞	116	앞	141	앞	166	뒤	191	앞	216	앞	241	앞
18	17	앞	42	앞	67	뒤	92	뒤	117	뒤	142	뒤	167	앞	192	앞	217	앞	242	뒤
19	18	뒤	43	뒤	68	뒤	93	앞	118	앞	143	앞	168	뒤	193	뒤	218	앞	243	뒤
20	19	뒤	44	앞	69	앞	94	뒤	119	앞	144	앞	169	뒤	194	뒤	219	뒤	244	뒤
21	20	앞	45	앞	70	뒤	95	뒤	120	앞	145	앞	170	뒤	195	앞	220	앞	245	뒤
22	21	뒤	46	앞	71	앞	96	뒤	121	뒤	146	앞	171	뒤	196	뒤	221	뒤	246	앞
23	22	뒤	47	앞	72	앞	97	뒤	122	앞	147	뒤	172	앞	197	뒤	222	뒤	247	앞
24	23	뒤	48	뒤	73	뒤	98	앞	123	앞	148	뒤	173	뒤	198	뒤	223	뒤	248	뒤
25	24	뒤	49	앞	74	뒤	99	뒤	124	앞	149	앞	174	앞	199	뒤	224	뒤	249	뒤
26	25	앞	50	앞	75	앞	100	앞	125	앞	150	뒤	175	뒤	200	뒤	225	뒤	250	앞

<그림 2>

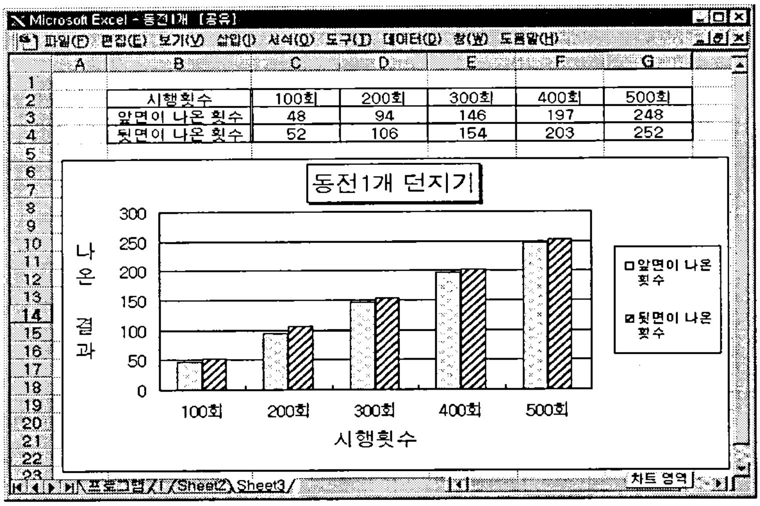


	A	C	D	F	G	I	J	L	M	O	P	R	S	U	V	X	Y	AA	AB	AD
28	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과	시행 횟수	나온 결과
29	251	앞	276	뒤	301	뒤	326	앞	351	뒤	376	앞	401	뒤	426	앞	451	앞	476	뒤
30	252	뒤	277	앞	302	앞	327	뒤	352	앞	377	뒤	402	앞	427	앞	452	앞	477	뒤
31	253	뒤	278	뒤	303	뒤	328	앞	353	뒤	378	앞	403	앞	428	뒤	453	뒤	478	앞
32	254	뒤	279	뒤	304	앞	329	앞	354	뒤	379	앞	404	뒤	429	뒤	454	앞	479	앞
33	255	앞	280	앞	305	뒤	330	뒤	355	앞	380	뒤	405	앞	430	앞	455	뒤	480	앞
34	256	앞	281	뒤	306	뒤	331	앞	356	앞	381	뒤	406	뒤	431	앞	456	뒤	481	뒤
35	257	뒤	282	뒤	307	뒤	332	앞	357	앞	382	앞	407	뒤	432	뒤	457	앞	482	뒤
36	258	뒤	283	뒤	308	뒤	333	뒤	358	뒤	383	앞	408	앞	433	앞	458	앞	483	뒤
37	259	뒤	284	앞	309	앞	334	뒤	359	앞	384	앞	409	앞	434	앞	459	앞	484	앞
38	260	뒤	285	앞	310	뒤	335	뒤	360	뒤	385	뒤	410	앞	435	뒤	460	뒤	485	뒤
39	261	뒤	286	앞	311	뒤	336	앞	361	뒤	386	뒤	411	뒤	436	앞	461	앞	486	뒤
40	262	뒤	287	앞	312	앞	337	앞	362	앞	387	뒤	412	뒤	437	앞	462	뒤	487	앞
41	263	앞	288	뒤	313	앞	338	앞	363	앞	388	뒤	413	뒤	438	뒤	463	뒤	488	앞
42	264	앞	289	앞	314	뒤	339	앞	364	앞	389	앞	414	뒤	439	앞	464	뒤	489	앞
43	265	뒤	290	앞	315	앞	340	앞	365	앞	390	앞	415	뒤	440	앞	465	뒤	490	앞
44	266	뒤	291	앞	316	뒤	341	앞	366	뒤	391	앞	416	앞	441	앞	466	뒤	491	뒤
45	267	뒤	292	뒤	317	뒤	342	뒤	367	뒤	392	앞	417	뒤	442	뒤	467	앞	492	앞
46	268	앞	293	앞	318	뒤	343	뒤	368	뒤	393	뒤	418	뒤	443	앞	468	뒤	493	앞
47	269	앞	294	뒤	319	뒤	344	뒤	369	뒤	394	뒤	419	뒤	444	앞	469	뒤	494	앞
48	270	뒤	295	앞	320	뒤	345	앞	370	앞	395	앞	420	앞	445	앞	470	뒤	495	뒤
49	271	뒤	296	앞	321	앞	346	앞	371	앞	396	앞	421	뒤	446	뒤	471	뒤	496	뒤
50	272	뒤	297	뒤	322	뒤	347	뒤	372	뒤	397	앞	422	뒤	447	앞	472	앞	497	뒤
51	273	앞	298	뒤	323	뒤	348	앞	373	뒤	398	뒤	423	뒤	448	앞	473	뒤	498	뒤
52	274	뒤	299	앞	324	앞	349	뒤	374	앞	399	뒤	424	뒤	449	뒤	474	뒤	499	뒤
53	275	뒤	300	뒤	325	뒤	350	앞	375	뒤	400	뒤	425	앞	450	앞	475	앞	500	앞

<그림 3>

이다. 교사가 프로젝터를 통해 이 프로그램을 학생들에게 제시할 수도 있으나 여건이 허락된 다면 학생들로 하여금 직접 실행해 보게 함으로써 흥미 유발과 탐구욕을 갖게 할 수 있다. 시행결과는 위와 같다(동전을 500회 던진 모의 실험임).

학생들은 위 실험에서 나온 결과를 한눈에 알아 볼 수 있게 정리해 보고 싶을 것이다. 동전의 앞면이 나온 횟수와 뒷면이 나온 횟수를 각각 시행횟수에 따라 비교해 보면 아래와 같다.



시행횟수	100회	200회	300회	400회	500회
앞면이 나온 횟수	48	94	146	197	248
뒷면이 나온 횟수	52	106	154	203	252

<그림 4>

각각 앞면과 뒷면이 나온 횟수를 조사할 때는 앞면의 경우 「=COUNTIF(시행횟수범위지정, “앞”)」,

뒷면의 경우 「=COUNTIF(시행횟수범위지정, “뒤”)」와 같은 통계함수를 사용하였고 그래프를 그릴 때는 먼저 위쪽의 표를 블록 설정한 후 챠트작성 도구 상자를 이용하여 세로 막대그래프를 선택하였다.

[문제] 위 표에서 앞면과 뒷면이 나온 횟수는 시행횟수와 비교하여 어떠한가?

[답] 앞면과 뒷면이 각각 시행횟수의 절반 정도로 나온다.

[문제] 위 세로막대그래프에서 시행횟수가 많아질수록 앞면과 뒷면이 나온 횟수를 나타내는 두 막대의 길이가 어떻게 되는가?

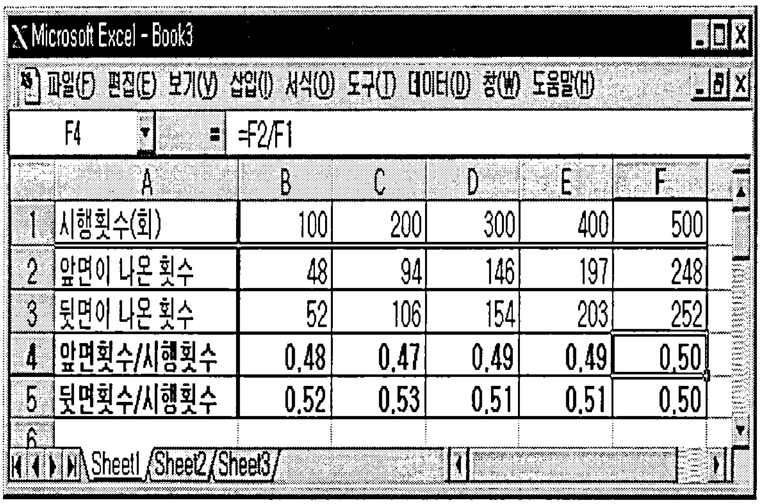
[답] 거의 두 막대의 길이가 같게 된다.

[문제] 시행횟수를 무한히 계속해 나간다면 앞면과 뒷면이 나올 가능성이 같다고 말할 수 있겠는가?

[답] 그렇다.

학생들로 하여금 시행 횟수를 마음대로 늘려가며 결과를 관찰 할 수 있는 기회를 갖게 하면 동전의 앞면과 뒷면이 나올 가능성이 같아짐을 인식하고 학생들은 수학적 확률을 계산하기 위한 전제조건인 ‘각 사건이 일어날 가능성은 같다’는 말의 의미를 이해하고 확률을 계산할 때 이러한 조건에 대해서 수긍할 수 있을 것이다.

이제 동전을 던졌을 때 앞면 또는 뒷면이 나오는 확률을 계산하기 위하여 위의 실험결과를 가지고 또 다른 표현으로 시행횟수에 따른 앞면과 뒷면의 빈도수의 비를 계산해 보게 하는 등 한 결과에 대한 다양한 분석 방법을 경험하게 한다.



	A	B	C	D	E	F
1	시행횟수(회)	100	200	300	400	500
2	앞면이 나온 횟수	48	94	146	197	248
3	뒷면이 나온 횟수	52	106	154	203	252
4	앞면횟수/시행횟수	0.48	0.47	0.49	0.49	0.50
5	뒷면횟수/시행횟수	0.52	0.53	0.51	0.51	0.50
6						

<그림 5>

[문제] 위의 표에서 시행횟수가 많아질수록 앞면이 나올 확률(=앞면횟수/시행횟수)은 어떻게 되는가?

[답] 0.48→0.47→0.49→0.49→0.50으로 약간의 편차는 있지만 0.5에 가깝게 된다.

[문제] 시행횟수가 많아질수록 뒷면이 나올 확률은 어떻게 되는가?

[답] 0.52→0.53→0.51→0.51→0.50 으로 0.5에 가깝게 된다.

[문제] 확률을 $\frac{(\text{사건 } A\text{가 일어날 경우의 수})}{(\text{일어날 수 있는 모든 경우의 수})}$ 로 계산한다면 한 개의 동전을 던질 때 앞면이 나올 확률은 얼마인가?

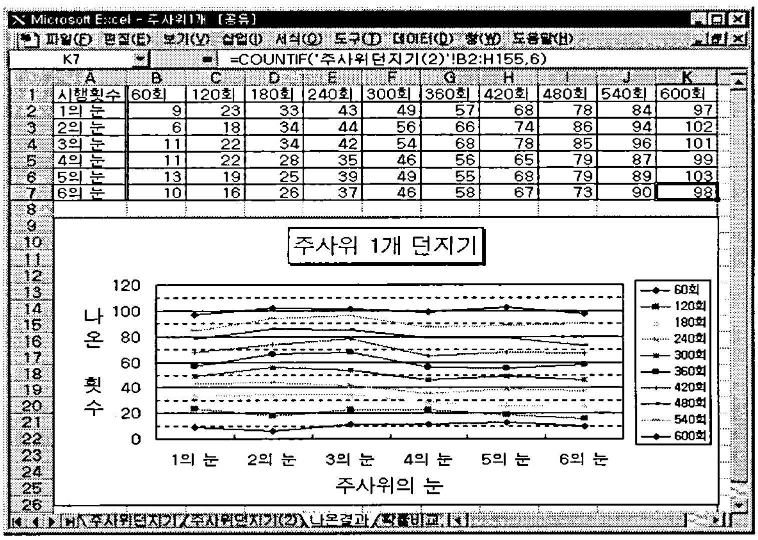
[답] 동전을 던지면 일어날 수 있는 모든 경우의 수는 앞면과 뒷면의 2가지이고 이중 앞면은 1가지이므로 구하는 확률은 1/2 즉 0.5 이다.

결국 위의 문제에 대한 답을 하면서 학생들은 직접 동전을 던져 보아서 얻은 통계적 확률(=0.5)과 수학적 확률을 계산하는 방법에 의해서 얻은 확률(= 1/2)이 같음을 알 수 있을 것이다. 즉 시행횟수가 충분히 클 때, 통계적 확률은 수학적 확률이 됨을 관찰할 수 있을 것이다. 이것은 '큰수의 법칙'으로 고등학교 교과서에 설명되어 있는데 중학교 교과과정에서도 충분히 설명되어질 수 있는 내용이다. 그러므로 수학적 확률을 이용하여 어떤 사건이 일어날 가능성을 예측할 수 있다는 것과 반대로 자연 현상이나 사회 현상과 같이 수학적 확률을 구하는 것이 곤란한 경우에는 통계적 확률을 이용할 수 있음을 주지시킬 수 있게 될 것이다.

[문제] 동전 2개를 던질 때 두 개 모두 뒷 면이 나오는 확률은 얼마인가?

[답] 동전 2개를 던지면 (앞, 앞), (앞,뒤), (뒤, 앞), (뒤, 뒤) 의 4가지 경우가 일어날 수 있고 그 중에서 모두 뒷면이 나오는 경우는 1가지이므로 구하는 확률은 1/4 이다.

이제 주사위 한 개를 던질 때도 동전을 던질 때와 마찬가지로 1부터 6까지의 눈이 나올 가능성이 모두 같은지 알아보기 위하여 모의실험을 하였다. 600회를 던진 후 나온 각 눈의 빈도수를 조사하여 이를 가지고 시행횟수에 따른 각 눈의 출현 횟수나타내는 또 다른 방법을 생각해 보게 하고, 표와 꺾은 선 그래프로 나타내어 본다.



시행횟수	60회	120회	180회	240회	300회	360회	420회	480회	540회	600회
1의 눈	9	23	33	43	49	57	68	78	84	97
2의 눈	6	18	34	44	56	66	74	86	94	102
3의 눈	11	22	34	42	54	68	78	85	96	101
4의 눈	11	22	28	35	46	56	65	79	87	99
5의 눈	13	19	25	39	49	55	68	79	89	103
6의 눈	10	16	26	37	46	58	67	73	90	98

<그림 6>

[문제] 주사위의 각 눈이 나오는 횟수가 시행 횟수에 따라 어떻게 다른가?

[답] 60회를 던졌을 때보다 시행횟수를 많이 하여 600회 던졌을 때가 각 눈의 출현 빈도수가 거의 비슷하게 나온다.

[문제] 각 눈이 나오는 횟수를 시행횟수와 비교하면 얼마나 되는가?

[답] 대부분 각 눈이 나오는 횟수가 시행횟수의 1/6정도로 나온다. 즉 주사위 1부터 6까지 눈의 수가 60회 던질 때는 10개 내외, 120회 던질 때는 20개 내외, 600회를 던졌을 때는 600/6 =100개 내외로 나온다.

[문제] 시행횟수를 무한히 계속하면 주사위의 각 눈이 나올 가능성은 같다고 할 수 있겠는가?

[답] 그렇다.

다시 시행결과에 대한 관찰을 위해 Excel의 간단한 기능을 이용하여 시행 횟수에 따른 각 눈의 출현 빈도수의 비를 계산해 보고 알아낸 성질을 발표하게 한다. 이때 각 학생들 마다 시행 결과는 다양하게 나타나지만 같은 결론에 이름을 학생 스스로 인지하도록 돕는다.

시행횟수	60회	120회	180회	240회	300회	360회	420회	480회	540회	600회
1의 눈이 나올 확률	0.15	0.19	0.18	0.18	0.16	0.16	0.16	0.16	0.16	0.16
2의 눈이 나올 확률	0.10	0.15	0.19	0.18	0.19	0.18	0.18	0.18	0.17	0.17
3의 눈이 나올 확률	0.18	0.18	0.19	0.18	0.18	0.19	0.19	0.18	0.18	0.17
4의 눈이 나올 확률	0.18	0.18	0.16	0.15	0.15	0.16	0.15	0.16	0.16	0.17
5의 눈이 나올 확률	0.22	0.16	0.14	0.16	0.16	0.15	0.16	0.16	0.16	0.17
6의 눈이 나올 확률	0.17	0.13	0.14	0.15	0.15	0.16	0.16	0.15	0.17	0.16

<그림 7>

[문제] 시행횟수가 많아질수록 각 눈이 나올 확률은 어떻게 되는가?

[답] 모두 0.17에 가깝게 된다.

[문제] 주사위 1개를 던질 때 각 눈이 나오는 수학적 확률은 무엇인가?

[답] 주사위의 눈이 모두 6개이고 각 눈은 1가지 씩이므로 1/6 (= 0.1666⋯ ≒ 0.17) 이다.

따라서 학생들은 주사위를 던질 때도 동전을 던질 때와 마찬가지로 수학적 확률을 이용하여 간단히 확률을 계산할 수 있음을 알게 될 것이다. 이것은 동전이나 주사위뿐만 아니라 다른 모든 경우에 대해서도 적용할 수 있을 것이다.

[문제] 주사위 2개를 던질 때 눈의 합이 4가 될 확률은 얼마인가?

[답] 눈의 합이 4가 되는 경우는 (1,3), (2,2), (3,1)의 3가지가 있으므로 확률은 $\frac{3}{36}=\frac{1}{12}$ 이다.

활용2. 이항분포와 정규분포와의 관계

성공할 확률이 p인 베르누이 시행을 n번 독립적으로 반복시행 했을 때 성공횟수 x를 확률변수 X가 취하는 값이라고 하면 확률변수 X의 확률함수가

$$f(x)=\binom{n}{x}p^x(1-p)^{n-x},\ x=0,1,2,.n$$

로 주어지는 확률분포를 이항분포라 한다.

또한 정규분포란 어떤 확률변수에 대응하는 확률들이 종 모양의 곡선(bell-shaped curve)의 면적과 같은 분포를 말하는데 이 때 곡선 밑의 모든 면적은 1이고 그 모양이 x축에 근접할 정도로 무한히 퍼져 있으며 평균 μ에 대하여 완전 대칭이다.

정규분포를 갖는 확률변수의 확률밀도함수는 $f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}$, $-\infty< x <\infty$

이며 평균과 분산은 각각 μ와 σ^2이다.

이항분포는 표본의 크기 n이 크게 되면 계산하기가 어려워진다. $n=20$이고 $p=0.5$일 때의 히스토그램을 그려 부드러운 곡선으로 연결시키면 정규분포 곡선에 가까운 확률분포를 얻을 수 있다. 즉, 이항분포를 근사시키는 곡선이 정규분포임을 알 수 있다.

이 정규분포의 평균과 분산은 그것에 근사시키는 평균과 분산, 즉 np와 $np(1-p)$이다. 그러나 정규분포를 이용한 근사계산은 $np>5$이고 $np(1-p)>5$ 일 때 만족한 만큼 n이 충분히 클 때 좋은 근사 값을 제공하여 주는 것이지 아무 때나 사용되는 것이 아닌 만큼 이를 사용할 때는 주의를 요한다.

[문제] 이산확률변수 X 가 이항분포 B(400, 1/5)를 따를 때, 확률 P(X≥76) 을 구하여 보아라.

위 문제에 답을 하기 위해서는 P(X=76) + P(X=77) + ⋯ + P(X=400)을 계산해야겠지만 사실 이것을 계산하기란 쉬운 일이 아니다. 고등학교 수학Ⅰ을 배운 사람이라면 다음과 같은 방법으로 해결할 수 있을 것이다.

[답] $E(X) = 400\times\frac{1}{5} = 80$, $V(X) = 400\times\frac{1}{5}\times\frac{4}{5} = 64$ 이므로

$P(X\geq 76)=P(Z=\frac{X-80}{8}\geq\frac{76-80}{8}=-0.5) = 0.6915$ 이다.

위와 같이 풀 수 있는 이유는 「이항분포 B(n,p) 를 따르는 확률변수 X는 n이 충분히 클 때, 정규분포 N(np,npq) (단 $q=1-p$)를 따른다」와 같은 이항분포와 정규분포와의 관계 때문이다. 이것도 Excel을 이용하여 시행횟수 n이 증가함에 따라 이항분포를 따르는 확률변수 X 의 그래프가 정말로 정규분포의 그래프 모양으로 변하는지 확인해 볼 수 있다. Excel 에는 이항분포확률을 구하는 「BINOMIST」라는 함수가 있어서 쉽게 그래프를 그릴 수 있기 때문에 학생들이 교사와 함께 직접 그려보는 것도 어려운 일이 아니다.

[예제] 다음은 확률변수 X가 이항분포 B(n, $\frac{1}{5}$) 을 따르고 n=10, 20, 30, 40, 50 일 때의 그래프를 그려본 것이다.

즉 ${}_{10}Cr\,(\frac{1}{5})^r\,(\frac{4}{5})^{10-r}$, r=1, 2, 3, …, 7

${}_{20}Cr\,(\frac{1}{5})^r\,(\frac{4}{5})^{20-r}$, r=1, 2, 3, …, 11

${}_{30}Cr\,(\frac{1}{5})^r\,(\frac{4}{5})^{30-r}$, r=1, 2, 3, …, 15

${}_{40}Cr\,(\frac{1}{5})^r\,(\frac{4}{5})^{40-r}$, r=1, 2, 3, …, 18

${}_{50}Cr\,(\frac{1}{5})^r\,(\frac{4}{5})^{50-r}$, r=1, 2, 3, …, 20

일 때 r에 대응하는 이항분포확률을 계산하여 각 순서쌍을 좌표평면 위에 표시하여 연결한 것이다.

그래프를 그리기 위해서는 먼저 시행횟수 n의 값에 따른 이항분포확률을 계산하는 표를 작성해야 하는데 이러한 표를 작성하기 위한 프로그램은 다음과 같다.

	A	B	C	D	E	F
1	성공할 횟수(r)	n=10	n=20	n=30	n=40	n=50
2	1	=BINOMDIST(A2,10,0.2,FALSE)	=BINOMDIST(A2,20,0.2,FALSE)	=BINOM	=BINON	=BINOMD
3	2	=BINOMDIST(A3,10,0.2,FALSE)	=BINOMDIST(A3,20,0.2,FALSE)	=BINOM	=BINON	=BINOMD
4	3	=BINOMDIST(A4,10,0.2,FALSE)	=BINOMDIST(A4,20,0.2,FALSE)	=BINOM	=BINON	=BINOMD
5	4	=BINOMDIST(A5,10,0.2,FALSE)	=BINOMDIST(A5,20,0.2,FALSE)	=BINOM	=BINON	=BINOMD
6	5	=BINOMDIST(A6,10,0.2,FALSE)	=BINOMDIST(A6,20,0.2,FALSE)	=BINOM	=BINON	=BINOMD
7	6	=BINOMDIST(A7,10,0.2,FALSE)	=BINOMDIST(A7,20,0.2,FALSE)	=BINOM	=BINON	=BINOMD
8	7	=BINOMDIST(A8,10,0.2,FALSE)	=BINOMDIST(A8,20,0.2,FALSE)	=BINOM	=BINON	=BINOMD
9	8		=BINOMDIST(A9,20,0.2,FALSE)	=BINOM	=BINON	=BINOMD
10	9		=BINOMDIST(A10,20,0.2,FALSE)	=BINOM	=BINON	=BINOMD
11	10		=BINOMDIST(A11,20,0.2,FALSE)	=BINOM	=BINON	=BINOMD
12	11		=BINOMDIST(A12,20,0.2,FALSE)	=BINOM	=BINON	=BINOMD
13	12			=BINOM	=BINON	=BINOMD
14	13			=BINOM	=BINON	=BINOMD
15	14			=BINOM	=BINON	=BINOMD
16	15			=BINOM	=BINON	=BINOMD
17	16				=BINON	=BINOMD
18	17				=BINON	=BINOMD
19	18				=BINON	=BINOMD
20	19					=BINOMD
21	20					=BINOMD

<그림 8>

여기서 쓰여진 Excel language에 대해 알아보면

· = BINOMIST(A2, 10, 0.2, FALSE) : n=10이고 r=1일 때의 이항분포확률 nCr $(\frac{1}{5})^1(\frac{4}{5})^9$을 구하기 위한 함수식이다. BINOMIST는 이항분포확률을 구할 때 쓰는 워크시트함수이고 A열에 r의 값이 1부터 20까지 들어 있으므로 A2에 있는 $r=1$일 때를 먼저 구한 것이고 10은 시행횟수 n의

값을 나타내며 0.2는 1회의 시행에서 사건이 일어날 확률이 $\frac{1}{5}$이므로 이것을 소수로 나타낸 것이고 FALSE는 확률밀도함수 (즉 P(X=r))를 구할 때 쓰는 것인데 이 부분이 TRUE이면 누적확률함수(즉 P(X≤r))를 구하는 것이 된다.

나머지 행은 2행을 블록지정을 한 후 마우스로 드래그하면(복사) 자동적으로 식이 채워진다.

얻어진 표는 다음과 같다(Excel 화면에서 ctrl+~ 키를 누르면 프로그램에 쓰여진 함수 식으로 화면이 표시되었다가 다시 누르면 결과수치로 화면이 전환된다). 즉 n번 실행해서 r번 성공할 확률들이 이항 분포를 가지고 계산되어 나온다.

Microsoft Excel - 이항분포확률

파일(F) 편집(E) 보기(V) 삽입(I) 서식(O) 도구(T) 데이터(D) 창(W) 도움말(H)

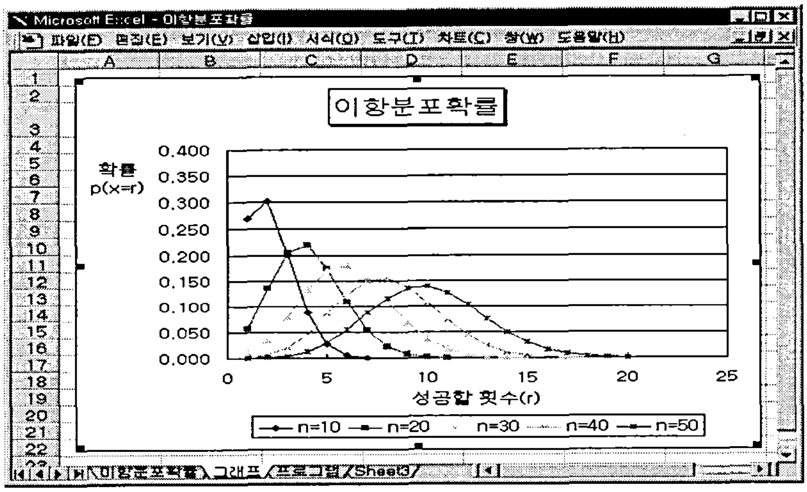
C3 =BINOMDIST(A3,20,0,2,FALSE)

	A	B	C	D	E	F
1	성공할 횟수 (r)	n=10	n=20	n=30	n=40	n=50
2	1	0.268	0.058	0.009	0.001	0.000
3	2	0.302	0.137	0.034	0.006	0.001
4	3	0.201	0.205	0.079	0.021	0.004
5	4	0.088	0.218	0.133	0.047	0.013
6	5	0.026	0.175	0.172	0.085	0.030
7	6	0.006	0.109	0.179	0.125	0.055
8	7	0.001	0.055	0.154	0.151	0.087
9	8		0.022	0.111	0.156	0.117
10	9		0.007	0.068	0.139	0.136
11	10		0.002	0.035	0.107	0.140
12	11		0.000	0.016	0.073	0.127
13	12			0.006	0.044	0.103
14	13			0.002	0.024	0.075
15	14			0.001	0.011	0.050
16	15			0.000	0.005	0.030
17	16				0.002	0.016
18	17				0.001	0.008
19	18				0.000	0.004
20	19					0.002
21	20					0.001

이항분포확률 / 그래프 / 프로그램 / Sheet3 /

<그림 9>

그래프를 얻기 위해서는 표에서 A2부터 F21까지를 블록 설정한 후에 차트 마법사를 눌러 차트 종류에서 분산형을 선택하고 그 외 적당한 옵션들을 선택하면 된다. 학생들로 하여금 시행횟수를 늘려가며 결과를 관찰하게 하고 그 결과들을 그래프로 나타내보게 한다. 시행횟수가 증가함에 따라 확률변수의 분포 형태가 종모양을 이룸을 알고 결론을 유도해 보게 한다.



<그림 10>

[문제] 위 이항분포확률의 그래프는 n의 값이 10에서 50까지 증가함에 따라 어떤 모양으로 변하는가?

[답] n의 값이 증가함에 따라 그 분포가 좌우 대칭인 종모양의 정규분포 곡선과 비슷하게 된다.

[문제] 위의 결과로 이항분포를 따르는 확률변수 X는 n이 충분히 클 때 (보통 30이상일 경우) 정규분포를 따른다고 할 수 있겠는가?

[답] 그렇다.

이상은 이산확률변수와 연속확률변수에 대해서 공부하는 고등학교 교과과정에서 설명될 수 있는 예일 것이다.

활용3. 중심 극한 정리

평균이 μ이고 분산이 σ^2인 모집단에서 추출한 크기 n인 확률표본 $X_1, X_2, \cdots, X_n$ 으로부터의 표본평균 X의 평균 E(X)는 모평균 μ와 같고 분산 Var(X)는 모분산 σ^2을 표본의 크기 n으로 나눈 값과 같다. 즉 $E(X)=\mu$, $Var(X) = \frac{\sigma^2}{n}$이다.

다시 말하면 X의 분산은 표본의 크기 n에 비례하여 감소하기 때문에 X의 표본분포는 표본의 크기가 증가할수록 모평균 μ 주위에 밀집하게 된다. 또 X 의 분포는 모집단의 형태에 따라 좌우되지만 임의의 모집단 분포에 대해서도 일반적으로 위와 같은 성질이 성립하게 된다. 따라서 정규모집단에서 추출한 표본들의

평균인 표본평균 X는 평균이 μ이고 분산이 $\frac{\sigma^2}{n}$인 정규분포를 따른다. 그런데 만약 모집단의 분포가 정규분포를 이루지 않는다면 일반적으로 표본평균 X의 분포가 정규분포를 따르지 않는다. 그

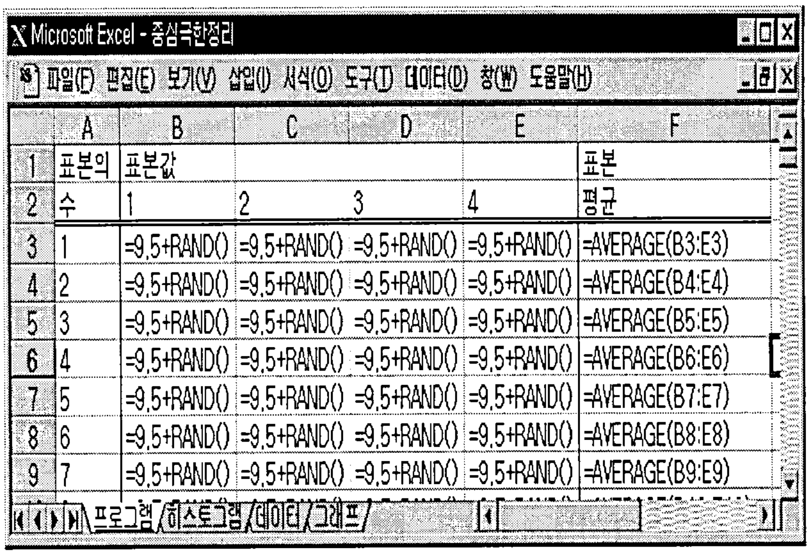
러나 표본의 크기 n이 충분히 크면 어떤 임의의 모집단에서의 확률표본이라 하더라도 표본평균 X의 분포를 정규분포로 근사시킬 수 있게 되는데 이것이 바로 중심 극한정리이며 통계학에서 가장 중요한 결과 중의 하나이다.

즉 평균이 μ이고 분산이 δ^2인 임의의 모집단으로부터 뽑은 확률표본의 평균의 분포는 표본의 크기가 크면 클수록 평균이 μ이고 분산이 $\frac{\delta^2}{n}$인 정규분포에 따른다.

다음은 중심 극한 정리(Central limit theorem)를 모의 실험을 통해 보이는 과정에 대해서 설명하였다.

[예제] 어느 공장에서 생산되는 제품들의 무게는 9.5g에서 10.5g 까지 균등하게 퍼져 있으며 이들 제품들의 평균 무게는 10g이다. 이 제품들 중에서 4개를 임의로 선택하여 그 길이를 잰 후 평균무게를 기록하였다. 500번 표본을 뽑은 후 표본평균의 분포를 그려보면 평균무게의 분포가 정규분포에 근사해 가는지 알아보자.

이 예는 모집단이 정규분포가 아님에도 표본들의 평균무게(X)가 표본의 수가 충분히 많아짐에 따라 정규분포를 따르는가 알아보는 문제이다. 이를 위한 프로그램은 다음과 같다



	A	B	C	D	E	F
1	표본의	표본값				표본
2	수	1	2	3	4	평균
3	1	=9.5+RAND()	=9.5+RAND()	=9.5+RAND()	=9.5+RAND()	=AVERAGE(B3:E3)
4	2	=9.5+RAND()	=9.5+RAND()	=9.5+RAND()	=9.5+RAND()	=AVERAGE(B4:E4)
5	3	=9.5+RAND()	=9.5+RAND()	=9.5+RAND()	=9.5+RAND()	=AVERAGE(B5:E5)
6	4	=9.5+RAND()	=9.5+RAND()	=9.5+RAND()	=9.5+RAND()	=AVERAGE(B6:E6)
7	5	=9.5+RAND()	=9.5+RAND()	=9.5+RAND()	=9.5+RAND()	=AVERAGE(B7:E7)
8	6	=9.5+RAND()	=9.5+RAND()	=9.5+RAND()	=9.5+RAND()	=AVERAGE(B8:E8)
9	7	=9.5+RAND()	=9.5+RAND()	=9.5+RAND()	=9.5+RAND()	=AVERAGE(B9:E9)

<그림 11>

· = 9.5+RAND() : 제품의 무게가 9.5g에서 10.5g 사이에 균등하게 퍼져 있으므로 표본값들도 이 사이에서 구해져야 한다. RAND() 함수는 0과 1사이의 임의의 난수를 생성하므로 9.5 에 생성된 난수를 더하면 9.5 와 10.5사이의 난수가 만들어진다.

· = AVERAGE(B3:E3) : 4개의 표본값들의 평균을 구하는 함수식이다.

500번 표본을 뽑기 때문에 3행의 식을 499개의 행에 복사해야 하는데 이 때는 마우스를 드래그함

으로써 간편하게 하면 된다.

아래는 표본평균들의 분포를 나타내는 그래프를 그리기 위한 도수분포표를 작성하기 위한 프로그램이다.

여기서 쓰인 FREQUENCY(F3:F502,H3:H23)는 빈도수를 구하기 위한 워크시트함수인데 F3에서 F502까지의 500개의 표본평균들 중에서 H3부터 H23까지 있는 계급에 해당되는 개수를 구하는 것이다.

얻어진 결과는 다음과 같다.

	H	I
2	계급	빈도수
3	9.5	=FREQUENCY(F3:F502,H3:H23)
4	9.55	=FREQUENCY(F3:F502,H3:H23)
5	9.6	=FREQUENCY(F3:F502,H3:H23)
6	9.65	=FREQUENCY(F3:F502,H3:H23)
7	9.7	=FREQUENCY(F3:F502,H3:H23)
8	9.75	=FREQUENCY(F3:F502,H3:H23)
9	9.8	=FREQUENCY(F3:F502,H3:H23)
10	9.85	=FREQUENCY(F3:F502,H3:H23)
11	9.9	=FREQUENCY(F3:F502,H3:H23)
12	9.95	=FREQUENCY(F3:F502,H3:H23)
13	10	=FREQUENCY(F3:F502,H3:H23)
14	10.05	=FREQUENCY(F3:F502,H3:H23)
15	10.1	=FREQUENCY(F3:F502,H3:H23)
16	10.15	=FREQUENCY(F3:F502,H3:H23)
17	10.2	=FREQUENCY(F3:F502,H3:H23)
18	10.25	=FREQUENCY(F3:F502,H3:H23)
19	10.3	=FREQUENCY(F3:F502,H3:H23)
20	10.35	=FREQUENCY(F3:F502,H3:H23)
21	10.4	=FREQUENCY(F3:F502,H3:H23)
22	10.45	=FREQUENCY(F3:F502,H3:H23)
23	10.5	=FREQUENCY(F3:F502,H3:H23)
24		

<그림 12>

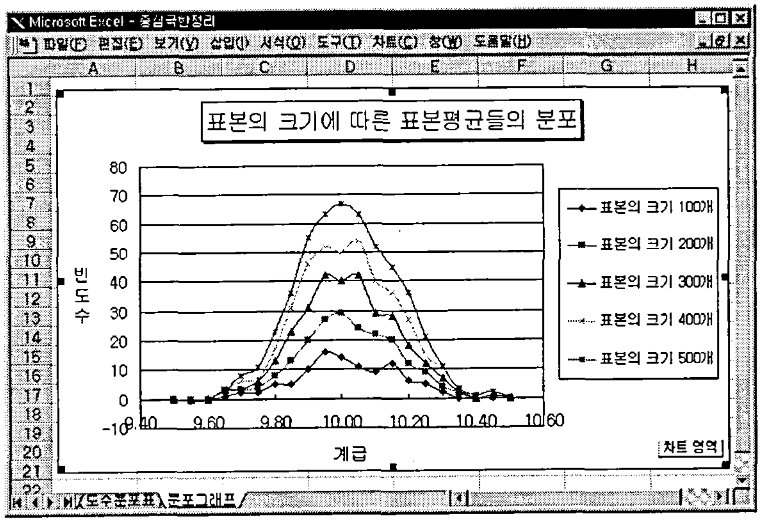
이제 표본평균들의 분포가 정규분포에 근사하는지 알아보기 위해서 그래프를 그려보면 되는데

	A	B	C	D	E	F	G
1	표본의	표본값				표본	
2	수	1	2	3	4	평균	
3	1	9.71	10.32	9.63	10.04	9.93	
4	2	9.72	9.85	9.62	10.08	9.82	
5	3	10.06	9.98	9.69	9.99	9.93	
6	4	10.00	10.03	9.51	9.92	9.87	
7	5	10.01	10.49	10.37	9.61	10.12	
8	6	9.69	10.28	9.88	10.35	10.05	
9	7	9.68	10.13	10.11	9.83	9.94	
10	8	9.98	10.22	9.75	9.96	9.98	
11	9	9.75	10.29	9.93	9.81	9.94	
12	10	10.18	9.97	9.89	10.44	10.12	
13	11	10.46	9.80	9.74	9.57	9.89	
14	12	9.81	10.12	9.95	9.80	9.92	
495	493	10.04	9.85	9.88	10.00	9.94	
496	494	10.37	9.82	10.44	10.19	10.20	
497	495	10.46	9.61	9.69	10.46	10.05	
498	496	9.70	10.18	9.97	10.29	10.04	
499	497	9.72	9.61	10.34	10.16	9.96	
500	498	9.70	10.25	9.75	9.75	9.86	
501	499	9.55	9.51	10.41	10.39	9.97	
502	500	10.48	9.54	10.07	9.52	9.90	

<그림 13>

이 때 위에서 작성된 도수분포표를 이용한다. 차트마법사를 누르고 분산형 그래프(직선연결형 분산형 그래프)를 선택한 후 데이터 범위는 빈도수에 있는 값들을 블록지정하고 x축 범위는 계급값을 블록지정한 후 확인을 누른다. 그 외 그래프의 모양을 꾸미고 싶으면 적당한 옵션들을 선택하면 된다.

다음 그래프를 보면 원래 모집단은 정규분포가 아니지만 상대적으로 큰 표본에서(위 문제에서는 500개의 표본평균들을 조사하였다) 표본평균들의 분포곡선이 정규분포의 그래프와 근사해 감을 알 수 있다.



<그림 14>

[문제] 앞의 예제에서 원래 모집단의 분포는 어떠하였나?

[답] 9.5와 10.5 사이에서 일양분포하는 균일모집단이었다. (즉 정규분포가 아니었다.)

[문제] 500개 표본평균들의 분포곡선의 모양은 어떠한가?

[답] 종 모양의 정규분포곡선과 비슷하다.

[문제] 위 모의실험의 결과로 알 수 있는 것은 무엇인가?

[답] 상대적으로 큰 표본에서 표본의 크기가 증가할수록 표본평균들의 분포는 모집단의 분포 형태에 상관없이 근사적으로 정규분포를 한다.

위의 모의실험들은 짧은 시간 내에 많은 결과들을 경험하게 함으로써 학생들이 파악하기 힘든 확률적 개념들을 받아 들이는 데 큰 도움이 되리라고 생각한다.

5. 결론

본 논문에서는 주로 확률·통계와 관련한 여러 가지 문제들을 해결하기 위한 모의실험도구로써

Excel의 활용방법에 대해서 설명하였다.

학습과정중의 중요한 정리나 여러 개념들 사이의 관계를 중·고등학교 학생들에게 논증적으로 증명하여 설명하기란 참 어려운 일이다. 그러나 확률 통계학습에 적절히 활용할 수 있는 여러 가지 모의실험을 제시함으로써 흥미와 탐구심을 유발시켜 확률 및 확률 분포의 개념 지도의 효율성을 높일 수 있을 것이다. 즉 이는 단순한 이론적 수치계산이 아닌 구체적 경험을 제시하여 학생들에게 확률적 상황에 내재된 확률적 정보의 의미를 파악하게 함으로써 확률의 개념에 대한 이해를 돕고 확률 통계단원에 대한 흥미를 유발케 하고자 하는 의도로 만들어진 프로그램이라면 효율적인 교수 학습을 위해 꼭 필요한 것이다.

본 논문에서 제시한 다양한 컴퓨터 모의실험은 학생들이 적극적으로 실험에 임하고 교사가 유연성을 가지고 수업을 이끌어 감으로써 정체되고 수동적인 수업이 아닌 학생들 스스로 결론에 다가가는 능동적 수업이 되어 확률적 개념 파악과 확률적 사고 능력 향상에 도움이 될 것이다.

그러나 컴퓨터를 이용한 수업이 만능은 아니므로 Excel을 활용하여 확률·통계를 지도할 때 고려해야할 점이 있음을 간과해서는 안된다.

첫째, 단편적인 지식보다는 관계성을 고려하고 실험적 경험을 통해 가정에서 결론을 얻는 활동이 다양하게 이루어지도록 지도해야 한다.

둘째, 학생들이 이론적인 기법과 시뮬레이션적인 기법의 차이와 각각의 장점을 이해하는 것이 중요하다. 더욱 중요한 것은 두 가지 접근 모두가 가치 있다는 것을 인식하도록 지도해야한다. 예를 들어, 이론적 모델의 적용을 통해 결과를 얻고, 시뮬레이션을 통해 그 결과의 타당성을 고려할 수 있음을 알게 한다.

셋째, 많은 실생활 현상으로부터 자료의 분포는 정규곡선에 근사하므로 모든 학생들은 그 그래프의 기하학적 성질에 친숙해야 하며 문제해결을 위해 그 곡선과 관련된 확률표나 컴퓨터 소프트웨어를 사용할 수 있어야 한다.

이제 칠판과 분필만을 가지고 수업하던 시대는 지나갔다. 수학의 개념이나 원리의 학습을 돕고 탐구적인 학습습관을 기르도록 하기 위해 여러 가지 보조매체들이 등장하고 있다. 이러한 맥락에서 구하기 쉽고 이용하기 편리하며 여러 기능들을 갖추고 있는 Excel을 이용하여 수업한다면 보다 효율적인 학습효과를 기대할 수 있을 것이다.

참 고 문 헌

구광조 외 공역 (1995). <u>수학교육과정과 평가의 새로운 방향</u>, 서울: 경문사.

안윤기 외 공저 (1998). <u>Excel 97 데이터 처리</u>, 자유아카데미.

우정호 (1998). <u>학교 수학의 교육적 기초</u>, 서울대학교 출판부.

우정호·이경화 (1996). 확률개념의 교수학적 변환에 관한 연구, <u>대한수학회 논문집</u> **6(1)**, pp.125-144.

송필원. 기초 확률론을 가르치는 데 Computer Simulation 사용의 효과에 관한 연구.

장경윤 (1997). 스프레드쉬트를 이용한 수학교육활동, 건국대학교 논문집.

Coburn P.; Kelman P.; Robert N.; Snyde T. & Weiner C. (1985), *Practical guide to computers in education*, Menlo Park, CA: Addison-wesley.

Dambolena, I.G. (1986), *Using simulation statistics courses, College Microcomputer* **4(4)**, pp.339-343.

Deane, E. (1984). Arganbright, Mathematical Applications of an Electronic Spreadsheet, Two-Year College Mathematics Journal.

National Council of Techers of Mathematics. (1998). *Principles and standards for School Mathema tics*; Disscussion Draft, NCTM.