1. 확률) 같은 것을 포함하는 순열, 중복조합의 사용가능여부

수학적 확률은 교과서에 다음과 같이 정의되어있다.

> 어떤 시행의 표본공간 S가 n개의 근원사건으로 이루어져 있고, 각 근원사건이 일어날 가능성이 모두 같은 정도로 기대된다고 하자.
>
> 이 때 사건 A가 r개의 근원사건으로 이루어져 있으면 사건 A가 일어날 확률 $\mathrm{P}(A)$를 수학적 확률이라 하고, $\mathrm{P}(A)=\dfrac{n(A)}{n(S)}=\dfrac{r}{n}$.

$\mathrm{P}(A)=\dfrac{n(A)}{n(S)}=\dfrac{r}{n}$을 모르고 확률 문제를 푸는 학생은 없다.
하지만, 우리는 과연 위에서 밑줄을 그은 각 근원사건이 일어날 가능성이 모두 같은 정도로 기대된다고 하자. 라는 부분의 이해가 확실히 되어있는가에 대한 물음을 해볼 필요가 있다.

두 문제를 풀어보자. (주사위의 각 면엔 하나의 숫자만 적혀있다.)

① 여섯 면에 $1, 1, 2, 3, 4, 5$가 적혀있는 주사위를 던질 때, 1이 나올 확률은?

② 여섯 면에 $1, 1, 2, 3, 4, 5$가 하나씩 적혀있는 주사위를 던졌을 때 나올 수 있는 숫자 중 하나를 임의로 뽑을 때, 1이 뽑힐 확률은?

문제 ①의 정답은 $\dfrac{2}{6}=\dfrac{1}{3}$이다. 직관적으로 생각해봐도 1이 적힌 면의 수가 다른 숫자가 적힌 면의 수보다 많으므로 $\dfrac{1}{6}$보단 높은 확률일 것이다.
문제 ②의 정답은 $\dfrac{1}{5}$이다. 주사위를 던졌을 때 나올 수 있는 숫자는 1부터 5까지의 자연수이고 그 중 1을 뽑을 확률은 다섯 개 중 하나.

두 문제의 표본공간 S 는 똑같이 $\{1, 2, 3, 4, 5\}$이나 정답과 풀이과정이 전혀 다르다. 그 이유에 대해 자세히 알아보자.

첫 번째 문제는 표본공간의 근원사건 중 하나인 1이 기대되는 정도가 타 원소에 비해 크기 때문에 단순히 $\mathrm{P}(A)=\dfrac{n(A)}{n(S)}=\dfrac{r}{n}=\dfrac{1}{5}$로 풀어내면 안된다.

두 번째 문제는 표본공간의 모든 근원사건이 기대되는 정도가 같으므로 $\mathrm{P}(A)=\dfrac{n(A)}{n(S)}=\dfrac{r}{n}=\dfrac{1}{5}$로 정답을 낼 수 있다.

즉, 위의 정의 그대로 '근원사건이 기대되는 정도가 같을 때' 일때 만 확률을 $\dfrac{r}{n}$로 정답을 낼 수 있는 것이다.

(cf. 이러한 차이는 문항 발문의 뉘앙스에서 알 수 있으므로, 문제를 읽을 때 제대로 파악하는 것이 좋다.)

결론적으로

[근원사건이 일어날 가능성이 모두 같은 정도로 기대되는 표본공간]을 잡는다면, $\mathrm{P}(A)=\dfrac{n(A)}{n(S)}=\dfrac{r}{n}$에서 n이나 r을 구할 때 도구의 구애를 받지 않는다. (같은 것을 포함하는 순열, 중복조합 등등)]

(예제) : 중복조합 사용
주머니에 0부터 3까지 적힌 공 4개 중 하나를 뽑아 숫자를 확인한 후 다시 집어넣는 행위를 세 번 시행할 때, 확인한 세 수의 합이 3일 확률은?

(바른 풀이)
나온 순서대로 세 수를 x, y, z라 할 때, 전체 경우의 수는 x, y, z가 각각 4가지의 숫자 중 하나를 선택하므로 4^3.

세 수의 합이 3이다. => $x+y+z=3$이다. => x, y, z가 3이하의 음이 아닌 정수이므로 $x+y+z=3$을 만족하는 (x, y, z)의 개수는 ${}_3\mathrm{H}_3$와 같다.

따라서 정답은 $\dfrac{{}_3\mathrm{H}_3}{64}$.

(틀린 풀이)
4개의 숫자 중 하나를 확인한 후 다시 집어넣으므로 같은 숫자를 중복해서 뽑아낼 수 있다. 따라서 세 번 확인한 세 수의 조합의 전체 경우의 수는 ${}_4\mathrm{H}_3$이다.
~~~~~~~

위 풀이가 틀린 이유는 ${}_4\mathrm{H}_3$가지의 원소로 구성되어 있는 표본공간의 근원사건의 기대되는 정도가 각각 다르기 때문이다.
(예를 들어 3이 세 번 확인 되는 기대정도와 $1, 2, 3$이 한번 씩 확인되는 기대정도가 다르다. $3, 3, 3$과 달리 $1, 2, 3$은 $2, 3, 1$, $3, 2, 1$등의 조합이 있으니까.)

물론 확률의 분자를 계산할 때 위의 기대되는 정도를 보정하기 위한 스킬을 사용해준다면 정답으로 향할 수는 있겠지만, 그걸 연습할 바에 바른 풀이로 푸는 연습을 하는 것이 한 6.02×10^{23}배 정도 좋다.

구체적인 문제로 확인해보자.

다음은 기대모의고사 가, 나형에 출제되었던 3점 문항이다.

(문제.1) : 같은 것을 포함한 순열 사용
10개의 문자 $\mathrm{E, Y, E, C, O, N, T, A, C, T}$를 배열하여 만들 수 있는 문자열 중 하나를 고를 때, 두 개의 E가 모두 이웃한 문자열을 고를 확률은? [3점]

(해설)
10개의 문자로 만들 수 있는 문자열을 우선 표본공간으로 다 모은 후, 그 표본공간에서 순수하게 하나를 고르는 것이므로 표본공간의 모든 근원사건이 기대되는 정도가 같다. 따라서 표본공간의 원소의 개수인 $\dfrac{10!}{2!2!2!}$를 E가 두 개 이웃한 문자열의 개수인 $\dfrac{9!}{2!2!}$에서 나눠주면 정답이 나온다.
~~~~~~~

2. 확률) 같은 것을 다른 것으로 봐라.

이번에는 다음과 같은 문제를 풀어보자.

(문제.2)
10개의 문자 $\mathrm{E_1, Y, E_2, C_1, O, N, T_1, A, C_2, T_2}$를 배열하여 만들 수 있는 문자열 중 하나를 고를 때, $\mathrm{E_1, E_2}$가 이웃한 문자열을 고를 확률은? [3점]

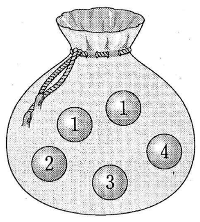
(풀이)
전체 경우의 수는 10!이다.
$\mathrm{E_1, E_2}$가 이웃하면 $\mathrm{E_1E_2}$ 혹은 $\mathrm{E_2E_1}$으로 이웃한다. $({}_2\mathrm{C}_1)$
이 덩어리를 하나의 원소로 보면 문제의 조건을 만족시키는 문자열의 개수는 이 덩어리와 나머지 8개의 문자로 구성된 9개의 문자+덩어리를 나열하는 경우의 수와 같으므로 ${}_2\mathrm{C}_1 \times 9!$ 이다. 여기에 10!을 나누면
문제.1의 정답과 똑같다.

문제.1에서 등장한 같은 문자들을 본질은 같지만 외형만 다른 문자들로 취급 (ex. E 두 개를 서로 다른 $\mathrm{E_1, E_2}$로 취급) 한 문제가 문제.2인데, 두 문제의 정답이 같은 것은 결코 우연이 아님을 배워가야한다.

확률문제에 같은 것이 포함된 것이 있다면, **본질만 같고 외형은 다른 문자들로 취급해서 문제를 풀도록 하자. 훨씬 편하다.**

대표 기출문제 (16학년도 9월 수학 B형 15번)

주머니에 1, 1, 2, 3, 4의 숫자가 하나씩 적혀 있는 5개의 공이 들어 있다. 이 주머니에서 임의로 4개의 공을 동시에 꺼내어 임의로 일렬로 나열하고, 나열된 순서대로 공에 적혀있는 수를 a, b, c, d라 할 때, $a \le b \le c \le d$일 확률은? [4점]



(풀이)
1, 1, 2, 3, 4을 1, 1′, 2, 3, 4로 취급하자.

4개를 꺼내어 $({}_5\mathrm{C}_4)$ 임의로 나열하자. (4!)

$i)$ a, b, c, d에 1, 1′이 모두 포함되어 있을 때
a와 b는 1, 1′ 이외의 수가 될 수 없다. (2!)
c와 d는 2, 3, 4 중 2개를 골라 큰 것을 d로 정해주면 된다. $({}_3\mathrm{C}_2)$

$ii)$ a, b, c, d에 1, 1′중 하나만 포함되어 있을 때,
$a = 1$ or $1'$ $({}_2\mathrm{C}_1)$
b, c, d는 반드시 2, 3, 4가 되어야 한다. (1)

따라서 정답은 $\dfrac{2! \times {}_3\mathrm{C}_2 + {}_2\mathrm{C}_1 \times 1}{{}_5\mathrm{C}_4 \times 4!} = \dfrac{1}{15}$ 이다.

3. 경우의 수) 케이스 분류의 이유

(문제)
일렬로 있는 6자리에 여자 A, B, C 와 남자 E, F, G를 앉힐 때, 여자가 2명 이상 이웃할 경우의 수는?

(해설)
$i)$ 여자가 2명 이웃할 때,
~~~~
$ii)$ 여자가 3명 모두 이웃할 때
~~~

해설에서 볼 수 있듯이, 위의 문제는
ㄱ.일렬로 있는 6자리에 여자가 2명 이웃할 경우의 수는?
ㄴ.일렬로 있는 6자리에 여자가 3명 이웃할 경우의 수는?
두 문제의 정답을 각각 구해서 더한 것과 같다.

위의 예시에서 왜 케이스를 분류해야 하는 이유을 알 수 있을 것이다.
그 이유는 **바로 추가조건을 획득함으로써 하나의 문제를 쉬운 여러 문제로 나눠서 풀 수 있기 때문**이다.

멋지게 한 번에 푸는 것도 좋다. 하지만 대부분의 문제는 한번에 풀기에는 너무 추상적일뿐더러 실수를 할 가능성도 많다.
따라서 케이스를 잘게 잘게 나눠서 기존의 문제를 쉬운 여러 문제로 나누어 구하는 연습을 꾸준히 할 수 있도록 하자.
(쪼갤 수 있는 최다의 케이스로 나눈 것이 바로 수형도이다.)

필자의 경우 대부분 다음의 기준으로 문제를 푸는 편이다.

$i)$ 문제의 조건의 중심이 되는 것
 ex. 여자가 이웃한다. => 여자가 어디서, 몇 명이나 이웃하는지에 대해 케이스를 나눈다.

$ii)$ 다른 것에 비해 원소의 수가 적은 것
 ex. 남자 2명과 여자 10명을 ~~~ => 남자 2명을 문제의 조건에 맞게 배열할 수 있는 케이스로 나눈 뒤 여자를 자유롭게 배치한다.

$ii)$와 같이 적은 것을 먼저 배치하는 것을 조삼모사라고 한다면 많은 것을 먼저 배치하는 것은 조사모삼이라고 할 수 있겠다.
무슨 말이냐?
적은 것을 먼저 배치하는 것은 풀이 초반엔 쉽지만 남은 ‘많은 것’들을 배열하기 어려울 수 있다. 하지만 ‘많은 것’을 먼저 배치하면 풀이 초반엔 어렵지만 많은 것을 이미 배치했으므로 남은 칸에 ‘적은 것’을 배치하는 것은 훨씬 간편하다는 말이다.

결론적으로, 그렇다면 케이스를 무엇을 기준으로 나누어야 하는가?
에 대한 100%짜리 정답은 없다.

경우의 수는 다양한 풀이가 가능하므로 **유형화 공부를 하는 것보다 본인만의 상황별 대처 알고리즘을 짜는 것**이 좋다. 여기서 상황별 대처 알고리즘이란 문제를 봤을 때 A를 先시도해보고, 안되면 B, C를 시도해보자. 라는 식의 알고리즘을 말한다.
이러한 알고리즘이 ‘케이스를 나누는 기준은 무엇인가?’에 대한 99%의 정답의 역할을 해줄 것이다.)

예시문제) 기대모의고사 6평 대비 가형 28번/나형 29번 (난이도 上)

0,1,2 중에서 중복을 허락하여 여섯 수를 택해 일렬로 배열하여 여섯 자리 자연수를 만들 때, 다음 조건을 만족시키는 자연수 N의 개수를 구하시오. [4점]

(가) 0은 2번 사용되었다. (나) 인접한 자리의 세 수의 합이 3이상이다. (예를 들어 201021은 $0+1+0=1$이므로 N에 포함되지 않는다.)

(hint : 케이스 나누는 기준 : 문제의 조건에 중심이 되는 것=0)
해설은 기대모의고사 해설지를 참고하세요.