Ⅲ. 추측통계(Inferential Statistics)를 위한 기본개념

1. 확률

【참고】 통계학에서 확률의 역할

〖예1〗 AB형인 남성과 O형인 여성이 만나 O형인 자녀를 낳을 수 있을까? A형인 자녀는?

〖예2〗 주사위 1개를 던져 6의 눈이 나올 가능성?

⇓

주사위 1개를 4번 던져 6의 눈이 1번 나올 가능성?

⇓

주사위 1개를 100번 던지면 6의 눈은? 이항분포

⇓

이항분포~정규분포... 예측, 추측 등에 활용

1-1. 경우의 수

1-1-1. 합의 법칙

① 두 사건 A, B가 동시에 일어나지 않을 때, A, B가 일어나는 경우의 수를 각각 m, n이라고 하면, 사건 A 또는 B가 일어나는 경우의 수는 $m+n$이다.

〖예〗 1에서 15까지의 자연수가 하나씩 적힌 공 중에서 하나를 뽑았을 때, 4의 배수 또는 5의 배수의 공이 나오는 경우의 수는?

[풀이]

1-1-2. 곱의 법칙

① 두 사건 A, B에 대하여 A, B가 일어나는 경우의 수를 각각 m, n이라고 하면, 두 사건 A, B가 잇달아 일어나는 경우의 수는 $m \times n$이다.

〖예〗 색깔이 각각 노란색, 분홍색, 흰색인 사탕 3개의 접시에서 두 개를 꺼내는 경우의 수는?

[풀이]

〚예〛 동전 3개를 던질 때 일어날 경우의 수는?

[풀이]

1-1-3. 순열의 수

① 서로 다른 n개에서 r개를 택하여 일렬로 배열하는 경우

$$ {}_{n}\mathrm{P}_{r} = \underbrace{n(n-1)(n-2)\cdots(n-r+1)}_{r\text{개}} $$

$(\text{단},\ 0 < r \leq n)$

② $$ {}_{n}\mathrm{P}_{r} = n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!} $$

③ ${}_{n}\mathrm{P}_{n} = n!$, $0! = 1$, ${}_{n}\mathrm{P}_{0} = 1$

〚예〛 50명 중에서 회장, 부회장, 총무를 임명하는 경우의 수는?

[풀이]

1-1-4. 중복순열의 수

① 서로 다른 n개에서 중복을 허용하여 r개를 택하는 순열

$$ {}_{n}\Pi_{r} = n^{r} $$

〚예〛 다섯 개의 문자 a, b, c, d, e에서 중복을 허용하여 3개를 택하여 일렬로 배열하는 경우의 수는?

[풀이]

1-1-5. 조합의 수

① 서로 다른 n개에서 순서를 생각하지 않고 r개를 택하는 경우

$$ {}_{n}\mathrm{C}_{r} = \frac{{}_{n}\mathrm{P}_{r}}{r!} = \frac{n(n-1)(n-2)\ \dots\ (n-r+1)}{r!} $$

$$ = \frac{n!}{r!(n-r)!}\ (\text{단},\ 0 < r \leq n) $$

〚예〛 집합 $\{a,\ b,\ c,\ d,\ e\}$의 부분집합 중에서 원소가 3개인 경우의 수는?

[풀이]

1-1-6. 중복조합의 수

① 서로 다른 n개의 원소에서 중복을 허락하여 r개를 택하는 경우의 수

$${}_n\mathrm{H}_r = {}_{n+r-1}\mathrm{C}_r$$

〖예〗 다섯 개의 문자 a, b, c, d, e에서 중복을 허용하여 3개를 택하는 경우의 수는?

[풀이]

1-1-7. 이항정리

① n이 자연수일 때, $(a+b)^n = {}_n\mathrm{C}_0 a^n + {}_n\mathrm{C}_1 a^{n-1} b^1 + \ldots + {}_n\mathrm{C}_n b^n$

② 이항계수 : ${}_n\mathrm{C}_0$, ${}_n\mathrm{C}_1$, $\cdots$, ${}_n\mathrm{C}_r$, $\cdots$, ${}_n\mathrm{C}_n$

③ ${}_n\mathrm{C}_r = {}_n\mathrm{C}_{n-r}$이므로 '$a^{n-r}b^r$의 계수 = $a^r b^{n-r}$의 계수' 임

〖예〗 $\left(x^2 - \dfrac{2}{x}\right)^3$의 전개식에서 x^3의 계수는?

[풀이]

1-2. 확률

1-2-1. 시행과 사건

① 시행 ; 같은 조건에서 여러 번 반복할 수 있고, 그 결과가 우연에 의하여 좌우되는 실험이나 관찰

② 사건 ; 어떤 시행에서 일어날 수 있는 모든 경우의 집합에 대하여 그 부분집합

〚예〛 한 개의 주사위를 던지는 행위(=시행)

이 시행에서 일어날 수 있는 모든 경우의 집합의 표본공간

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

A사건 '2의 눈이 나온다.' ; $A=\{2\}$

B사건 '홀수의 눈이 나온다.' ; $B=\{1,\ 3,\ 5\}$

③ 두 사건 A, B에 대하여

㉠ 'A 또는 B가 일어나는 사건' ; 합사건 $A\cup B$

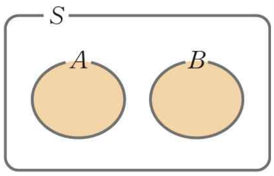
㉡ 'A와 B가 동시에 일어나는 사건' ; 곱사건 $A\cap B$

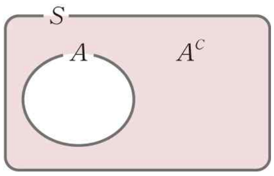
㉢ A와 B가 동시에 일어나지 않을 때, 곧 $A\cap B=\varnothing$인 사건

; A와 B는 배반사건

㉣ 어떤 사건 A에 대하여 A가 일어나지 않는 사건

; A의 여사건, A^C





1-2-2. 확률

① 확률 ; 시행에서 어떤 사건이 일어날 가능성을 수치로 나타낸 것

$$\mathrm{P}(A)$$

② 수학적 확률

㉠ 어떤 시행에서 일어날 수 있는 모든 경우의 수 n,

㉡ 각 경우가 일어날 가능성이 모두 같은 정도로 기대된다고 할 때, 사건 A가 일어나는 경우의 수 r

㉢ 사건 A가 일어날 확률 $\mathrm{P}(A)=\dfrac{r}{n}$

〚예〛 한 개의 주사위, 짝수가 나올 확률?

[풀이]

한 개의 주사위를 던지는 시행, 모든 경우 1, 2, 3, 4, 5, 6의 눈이 나오는 가능성이 모두 동일 사건 A가 일어날 확률 $\mathrm{P}(A)=\dfrac{3\text{가지}}{6\text{가지}}=\dfrac{1}{2}$

③ 통계적 확률

㉠ 같은 시행을 n번 반복

㉡ 사건 A가 일어난 횟수를 r_n

㉢ n이 충분히 커짐에 따른 상대도수 $\dfrac{r_n}{n}$의 극한값 p

〖예〗 2007년의 주요 자동차 생산국의 자동차 생산량

국가	일본	미국	중국	독일	한국	프랑스	브라질	스페인	캐나다	인도	합계
생산대수(대)	1159만 6000	1075만 1000	882만 2000	619만 5000	408만 6000	309만 1000	297만 2000	288만 9000	257만 8000	224만 5000	5522만 5000

[풀이] 통계적 확은 $\frac{4086000}{55225000} \fallingdotseq 0.074$

㉣ 실제로 통계적 확률을 구할 때에는 n의 값을 한없이 크게 할 수 없으므로 n이 충분히 클 때의 상대도수 $\frac{r_n}{n}$으로 통계적 확률을 대신함

㉤ 사건 A가 일어나는 상대도수 $\frac{r_n}{n}$은 시행 횟수 n을 충분히 크게 하면 할수록 수학적 확률에 가까워지는 경향이 있는 것으로 알려져 있음

1-2-3. 확률의 기본 성질

① 임의의 사건 A에 대하여 $0 \le \mathrm{P}(A) \le 1$
② 일어날 수 있는 모든 경우의 집합 S에 대하여 $\mathrm{P}(S) = 1$
③ 절대로 일어날 수 없는 사건 $\varnothing$에 대하여 $\mathrm{P}(\varnothing) = 0$
* 반드시 일어날 사건 = 전사건
절대로 일어날 수 없는 사건 = 공사건

1-2-4. 확률의 덧셈정리

① 두 사건 A, B에 대하여 사건 A 또는 B가 일어나는 확률 $\mathrm{P}(A \cup B)$는

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$$

② 특히, A, B가 배반사건이면

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$$

③ n개의 사건 A_1, A_2, $\cdots$, A_n이 서로 배반사건이면

$$\mathrm{P}(A_1 \cup A_2 \cup \cdots \cup A_n) = \mathrm{P}(A_1) + \mathrm{P}(A_2) + \cdots + \mathrm{P}(A_n)$$

〖예〗 필통의 빨간색 색연필 6자루, 파란색 색연필 4자루에서 임의로 두 자루를 꺼낼 때, 모두 같은 색일 확률?

[풀이]

1-2-5. 여사건의 확률

① 사건 A의 여사건 A^C의 확률은

$$\mathrm{P}(A^C) = 1 - \mathrm{P}(A)$$

〖예〗 6개의 통곡물 과자와 4개의 초콜릿 칩 과자 중 임의로 3개를 고를 때, 초콜릿 칩 과자가 적어도 한 개 포함될 확률은?

[풀이]

1-3 조건부확률과 곱셈정리

1-3-1. 조건부확률

① 사건 A가 일어났을 때 사건 B가 일어날 확률

② $\mathrm{P}(B|A)$

> 사건 A가 일어났을 때의 사건 B의 조건부확률
>
> $$\mathrm{P}(B|A)=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(A)}\ \ (\text{단, } \mathrm{P}(A)>0)$$

〖예〗 한 개의 주사위를 던져서 홀수의 눈이 나왔을 때(A), 그것이 소수(B)일 확률

[풀이]

$A=\{1,3,5\}\ B=\{2,3,5\},\ A\cap B=\{3,5\}$

$\mathrm{P}(A)=\frac{3}{6},\ \mathrm{P}(B)=\frac{3}{6},\ \mathrm{P}(A\cap B)=\frac{2}{6}$

$$\mathrm{P}(B|A)=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(A)}=\frac{2}{6}\div\frac{3}{6}=\frac{2}{3}$$

	$n(A)$	$n(A^C)$	합계
$n(B)$	2	1	3
$n(B^C)$	1	2	3
합계	3	3	6

$p=\frac{2}{3}$

1-3-2. 확률의 곱셈정리

① $\mathrm{P}(A)>0,\ \mathrm{P}(B)>0$일 때 조건부확률 $\mathrm{P}(B|A),\ \mathrm{P}(A|B)$는

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

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

> $\mathrm{P}(A)>0,\ \mathrm{P}(B)>0$일 때, 사건 $A\cap B$가 일어날 확률
>
> $$\mathrm{P}(A\cap B)=\mathrm{P}(A)\mathrm{P}(B|A)=\mathrm{P}(B)\mathrm{P}(A|B)$$

② 독립사건의 곱셈정리

㉠ 사건 B가 사건 A에 대하여 독립이면 $\mathrm{P}(B|A)=\mathrm{P}(B)$이다.

㉡ 사건 A도 사건 B에 대하여 독립이다.

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

$$=\frac{\mathrm{P}(A)\mathrm{P}(B)}{\mathrm{P}(B)}=\mathrm{P}(A)$$

㉢ 두 사건 A와 B는 서로 독립일 때, 독립사건이라 한다.

> 두 사건 A, B가 서로 독립이기 위한 필요충분조건은
>
> $$\mathrm{P}(A\cap B)=\mathrm{P}(A)\mathrm{P}(B)\ \ (\text{단,}\mathrm{P}(A)>0,\ \mathrm{P}(B)>0)$$

㉣ 두 사건이 서로 독립이 아닐 때 ; 종속사건

〖예〗 한 개의 주사위를 던질 때, 짝수의 눈이 나오는 사건을 A, 소수의 눈이 나오는 사건을 B, 5 이상의 눈이 나오는 사건을 C에서, 다음 두 사건이 서로 독립인지 종속인지 판별하여라.

① A와 B

② A와 C

③ B와 C

[풀이] A와 B는 종속사건이고, A와 C 그리고 B와 C는 독립사건이다.

【참고】 두 사건 A와 B가 서로 독립임을 보이려면

㉠ $\mathrm{P}(B) = \mathrm{P}(B|A) = \mathrm{P}(B|A^C)$를 밝히거나

㉡ $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$를 밝히면 된다.

1-3-3. 독립시행

① 어떤 시행을 되풀이할 때 각 시행의 결과가 그 다음의 시행의 결과에 아무런 영향을 주지 않는 시행

> 1회의 시행에서 사건 A가 일어날 확률이 p일 때, n회의 독립시행에서 사건 A가 r회 일어날 확률은
> ${}_n\mathrm{C}_r p^r (1-p)^{n-r}$ (단, $r = 0,\ 1,\ 2,\ \cdots\cdots,\ n$)

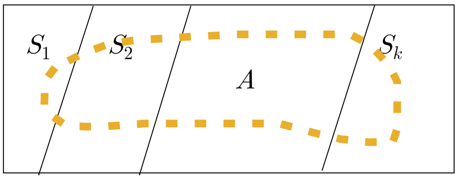
〖예〗 한 개의 주사위를 여러 번 던지는 시행

〖예〗 A 선수의 자유투 성공률 0.8, 3번의 자유투를 던질 때, 2번 성공할 확률은?

[풀이]

$${}_3\mathrm{C}_2(0.8)^2(0.2)^1 = 3 \times 0.64 \times 0.2 = 0.384$$

1-4. 베이즈공식



① 표본공간(S)이 서로 배반인 부분집합 S_1, S_2, $\cdots$, S_k에 의하여 분할되었다하고 A를 임의의 사건이라 할 때, 사건 A가 일어났다는 조건하에 S_j의 조건부 확률

$$P(S_j|A) = \frac{P(S_j \cap A)}{P(A)} = \frac{P(S_j)P(A|S_j)}{\sum_{i=1}^{k} P(S_i)P(A|S_i)}$$

② A의 전 확률의 공식(law of total probability)

; 표본공간(S)이 서로 배반인 부분집합 S_1, S_2, $\cdots$, S_k에 의하여 분할

㉠ $A = (A\cap S_1) \cup (A \cap S_2) \cup \cdots (A \cap S_k)$

㉡ A의 확률

$$\begin{aligned} P(A) &= P(A\cap S_1) + P(A\cap S_2) + \cdots + P(A\cap S_k) \\ &= P(S_1)P(A|S_1) + P(S_2)P(A|S_2) + \cdots + P(S_k)P(A|S_k) \\ &= \sum_{i=1}^{k} P(S_i)P(A|S_i) \end{aligned}$$

㉢ 다음 색맹실험표에서 색맹일 사건 A의 확률은?

	남자(B)	여자(B^C)	합계
색맹(A)	0.04	0.002	0.042
정상(A^C)	0.47	0.488	0.958
합계	0.51	0.490	1.000

[풀이] 색맹일 사건 A는 서로 배반인 두 사건 $A\cap B$와 $A\cap B^C$로 이루어지므로

$$A = (A\cap B) \cup (A\cap B^C)$$

이다. 따라서 $P(A) = 0.04 + 0.002 = 0.042$이다.

[예] 세 공장 A, B, C에서 각각 40%, 30%, 30%의 비율로 제품을 생산한다. 그리고 각 공장에서 불량품이 제조될 가능성이 각각 2%, 3%, 5%라 한다. 어떤 제품 하나를 임의로 선정했을 때, 이 제품이 불량품일 확률은? 임의로 선정된 제품이 불량품일 때, 그 제품이 A 공장에서 만들어졌을 확률은?

[풀이]