김은영

제4장 이항분포와 가설검정에서의 응용

4.1 서론

❑ 실험의 결과가 두 가지만 있는 경우(성공과 실패, 양호와 불량, 찬성과 반대)와 관련된 분포를 알아보려 한다. 이런 실험을 일정하게 반복할 때에 확률변수를 한 가지 가능한 경우의 도수로 정의할 수 있다.
❑ 일반적으로 모비율(population proportion)이 알려져 있다면 확률변수 X의 확률분포를 구할 수 있다.
❑ 그러나 모비율은 대개 알려져 있지 않다. 이 경우에도 모비율을 모수(parameter)로 하여 확률변수 X의 확률모형을 세울 수 있다.

4.2 베르누이 시행(Bernoulli trials)

(1) 베르누이 시행의 정의

각 시행은 두 가지 경우[성공(S)과 실패(F)] 중 한 가지에 해당한다.	
각 시행이 성공할 확률은 $P(S)=p$이고	단, $p+q=1$
실패할 확률은 $P(F)=q$이다.	
각 시행들은 독립적이다.	

(2) 베르누이 시행의 예: 한 개의 동전을 던지는 실험
① 가능한 결과: H, T
② 동전이 공정하게 만들어졌다면, $P(H)=P(T)=1/2$

(3) 모집단으로부터의 가능한 두 가지 표본추출법
① 복원표본추출법(sampling with replacement)
② 비복원표본추출법(sampling without replacement)

[예제] 상자에서 카드 꺼내기

상자 안에 15장의 카드가 있다. 5장의 카드는 빨간색이고 10장의 카드는 파란색이다. 상자에서 한 장의 카드를 꺼낼 때, 그 카드가 빨간색일 확률은?

① 복원표본추출법 (sampling with replacement)

사상 R: 빨간색 카드를 꺼내는 경우

$P(R) = 5/15$

② 비복원표본추출법 (sampling without replacement)

첫 번째 시도에서 빨간색 카드를 꺼낼 확률: $P(R) = 5/15$.

두 번째 시도에서 빨간색 카드를 꺼낼 확률: $P(R) = 4/14$.

4.3 이항분포(binomial distribution)

(1) 이항분포(binomial distribution)

<table>
<tr><td>n: 베르누이 시행의 횟수</td><td rowspan="3">확률변수 X: 이항확률변수
확률변수 X의 분포: 이항분포</td></tr>
<tr><td>p: 성공확률</td></tr>
<tr><td>X: n번 시행에서 성공한 횟수</td></tr>
<tr><td colspan="2">$f(x) = P[X = x] = \binom{n}{x} p^x q^{n-x},\ x = 0, 1, \cdots, n$</td></tr>
<tr><td colspan="2">▸ 평균(mean)=np
▸ 분산(variance)=npq 여기서, $q = 1 - p$
▸ 표준편차(standard deviation)= $\sqrt{npq}$</td></tr>
</table>

(2) 누적이항분포표(cumulative binomial distribution table), 부록 [표 1]

(n, p)가 주어지면 표에서 각 c에 해당하는 값은 아래 누적확률을 나타낸다.

$$P[X \le c] = \sum_{x=0}^{c} f(x)$$

이항분포

x	$f(x)$
0	$f(0)$
1	$f(1)$
2	$f(2)$
$\vdots$	$\vdots$
n	$f(n)$
합계	1

누적이항분포표

c	$P[X \le c] = \sum_{x=0}^{c} f(x)$
0	$f(0)$
1	$f(0)+f(1)$
2	$f(0)+f(1)+f(2)$
$\vdots$	$\vdots$
n	$f(0)+f(1)+f(2)+\cdots+f(n)=1$

[예제 4.4] 근육통을 치료하는 새로운 방법이 50%의 성공률을 거두는 것으로 알려져 있다. 15명의 환자를 대상으로 치료했을 때 다음의 확률은 얼마인가?

① 기껏해야 6명이 치료될 것이다.

② 6명 이상 10명 이하가 치료될 것이다.

③ 12명 이상 치료될 것이다.

【풀이】

확률변수 X: 치료된 환자의 수, $n=15$이고 $p=0.5$인 이항분포를 따른다.

① $P[X \le 6]=0.304$

② $P[6 \le X \le 10]=f(6)+f(7)+f(8)+f(9)+f(10)$

$$=P[X \le 10]-P[X \le 5]=0.941-0.151=0.790$$

③ $P[X \ge 12]=1-P[X \le 11]=1-0.982=0.018$

[예제 4.5] $n=3$이고 $p=0.5$인 이항분포의 평균과 표준편차를 구하라.

【풀이】

☺ 평균$=np=3\times 0.5=1.5$

☺ 표준편차$=\sqrt{npq}=\sqrt{3\times 0.5\times 0.5}=\sqrt{0.75}=0.866$