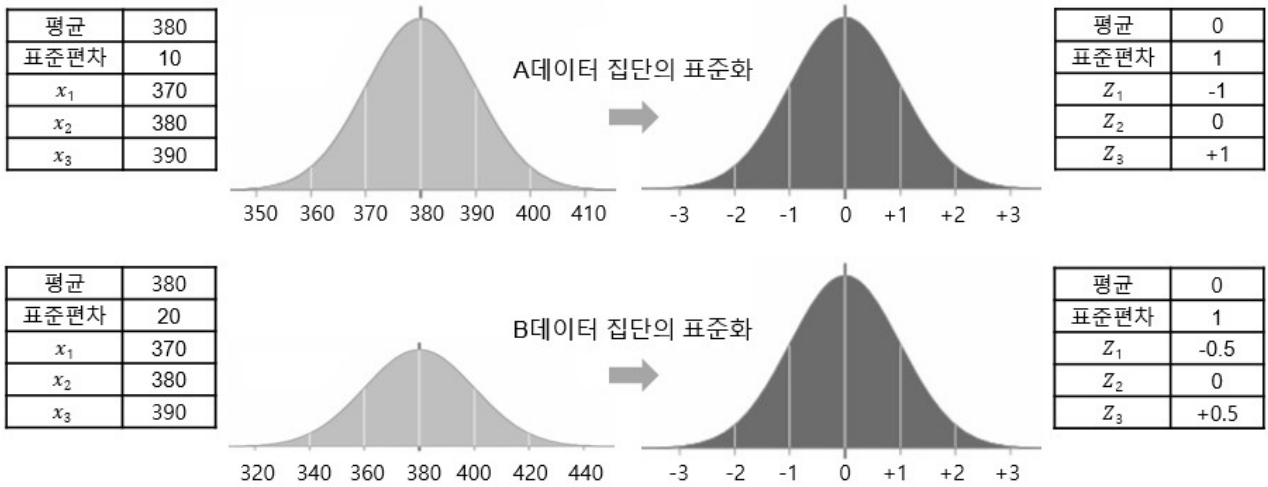


표준정규분포

표준정규분포(Standard Normal Distribution)는 정규분포를 표준화(Standardization)하여 얻어진 확률분포를 칭한다. 여기서 표준화란 정규분포의 변수 대신 아래의 식으로 정의되는 표준화변수 Z에 대한 확률분포를 얻는 것이다. 따라서 표준정규분포를 Z Distribution이라 칭하기도 한다.

$$Z = \frac{x - \mu}{\sigma} \quad (\text{여기서 } x \text{는 변수, } \mu \text{는 평균, } \sigma \text{는 표준편차})$$

정규분포를 표준화하여 얻어진 표준정규분포의 평균은 항상 0이고 표준편차 또한 항상 1이다. 표준화는 평균이 같지만 분산이 다른 즉, 표준편차가 다른 두 데이터 집단의 해석을 용이하게 한다. 다음은 변수의 평균이 380로 같고 표준편차가 각각 10, 20인 A, B 두 데이터 집단의 표준화 예이다.



예를 들어, 변수 400에 해당하는 Z값(Z Score 또는 Z Value라 칭함)이 A데이터 집단에서는 +2로 해당 변수가 평균으로부터 표준편차의 2배만큼 떨어진 곳에 위치하지만, B데이터 집단에서는 Z값이 +1이므로 B데이터 집단이 더 우수함을 알 수 있다.

표준화는 평균과 표준편차가 모두 다른 두 데이터 집단의 상호 해석을 용이하게 한다.

예를 들어 작은 키 순으로 번호를 받는 A, B, C 세 학급 학생들의 평균키가 168, 170, 172cm 이며, 표준편차가 각각 8, 5, 6 인 정규분포를 따른다고 하자. 만일 각 반에서 임의로 뽑힌 세 학생의 키가 모두 각 학급의 평균보다 큰 176, 177, 180cm일 때 이 세 학생 중 자신의 반에서 가장 높은 번호를 받을 가능성이 높은 학생은 표준화변수가 1.4로 가장 큰, 키 177cm인 학생임을 쉽게 확인할 수 있다.

Z Table(Standard Normal Table 또는 Unit Normal Table)은 지정된 범위 내에 존재하는 표준화변수 Z가 가지는 확률을 기록한 표이다. 아래의 표는 해당 표준화변수를 기준으로 그와 같거나 작은 변수들이 가지는 확률을 소수점 5자리까지 기록한 Z Table의 일부이다.

Z	+0.00	+0.01	+0.02	+0.03	+0.04	+0.05	+0.06	+0.07	+0.08	+0.09
0.1	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.50399	0.50399	0.50399
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524

만일 표준화변수가 0.77이라면, 이와 같거나 작은 변수가 가지는 확률 즉, $P(0.77 \leq Z)$ 는 위의 표를 통해 0.77935임을 쉽게 알 수 있다. 표준정규분포는 평균을 중심으로 좌우가 대칭이며, 모든 변수가 나타날 확률은 항상 1이므로 표준화변수 0.77보다 큰 변수들이 가지는 확률 즉, $P(Z > 0.77)$ 는 0.22065가 된다.