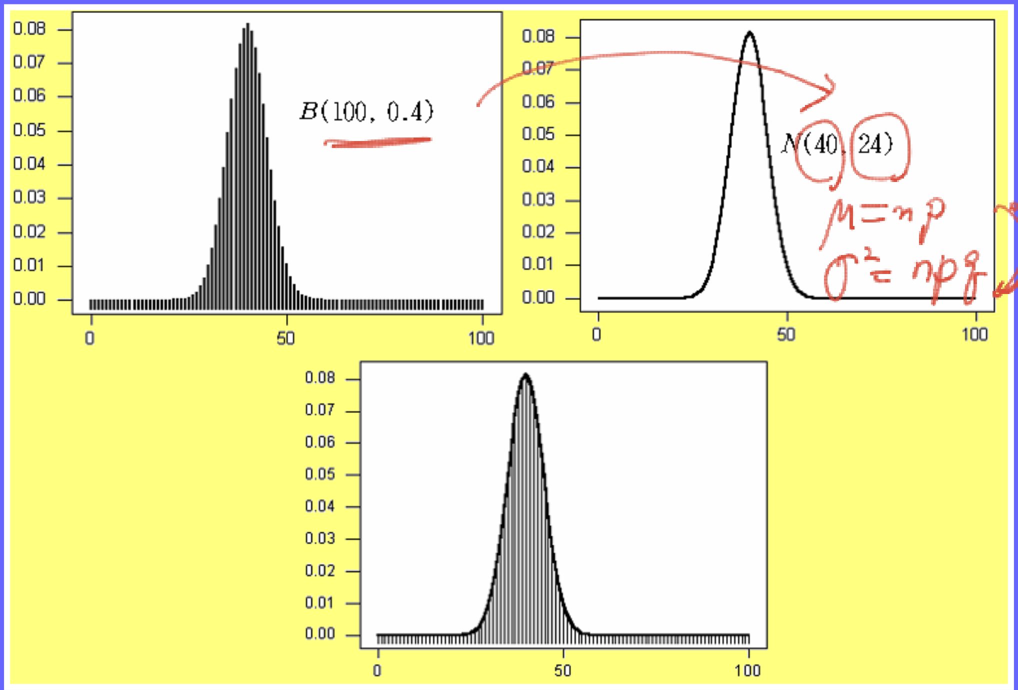
이항분포의 정규근사

● 모수 n과 p인 이항분포에 대하여 $np \geq 5$, $nq \geq 5$인 경우, n이 커질수록 이항분포는 평균 $\mu = np$, 분산 $\sigma^2 = npq$인 정규분포에 근사하는 것으로 알려져 있다. 따라서 $X \sim B(n, p)$, $np \geq 5$, $nq \geq 5$이면 X는 정규분포 $X \approx N(np, npq)$에 근사하며, 이것을 **이항분포의 정규근사**(normal approximation)라 한다.



● 이항분포의 근사확률 : $a, b = 0, 1, 2, \ldots, n$, $a < b$에 대하여 정규분포를 이용하여 다음과 같이 근사확률을 구할 수 있다.

$$P(a \le X \le b) \approx P\left(\frac{a-np}{\sqrt{npq}} \le Z \le \frac{b-np}{\sqrt{npq}}\right)$$

$X \sim B(n, p)$ $\qquad z_l \qquad z_u$

$$\boxed{X \approx N(np, npq)} \Rightarrow Z = \frac{X-\mu}{\sigma} = \frac{X-np}{\sqrt{npq}}$$

[예제 6] n=15 → p=0.4

$X \sim B(15, 0.4)$일 때 다음을 구하라.

(1) 이항분포표를 이용하여 $P(7 \le X \le 9)$를 구하라.

(2) 정규근사에 의해 $P(7 \le X \le 9)$를 구하라.

(3) 정규근사에 의해 $P(6.5 \le X \le 9.5)$를 구하라.

풀이

(1) $P(7 \le X \le 9) = P(X \le 9) - P(X \le 6) = 0.9662 - 0.6098 = 0.3564$

(2) $np = 15(0.4) = 6$이고 $npq = 15(0.4)(0.6) = 3.6$이므로 근사적으로 $X \approx N(6, 3.6)$ 이다. 따라서 7과 9를 표준화한다. :

$$z_l = (7-6)/\sqrt{3.6} = 0.53, \quad z_u = (9-6)/\sqrt{3.6} = 1.58$$

$$P(7 \le X \le 9) \approx P(0.53 \le Z \le 1.58) = P(Z \le 1.58) - P(Z \le 0.53) = 0.9429 - 0.7012 = 0.241$$

(3) 6.5와 9.5를 표준화한다. : $z_l = (6.5-6)/\sqrt{3.6} = 0.26, \quad z_u = (9.5-6)/\sqrt{3.6} = 1.84$

$$P(6.5 \le X \le 9.5) \approx P(0.26 \le Z \le 1.84) = P(Z \le 1.84) - P(Z \le 0.26)$$
$$= 0.9671 - 0.6026 = 0.3645$$

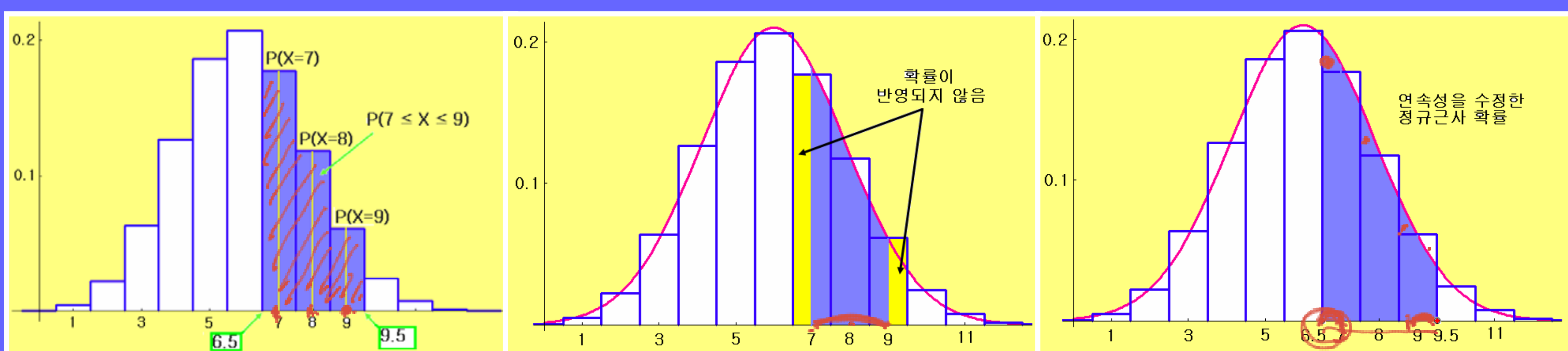
연속성 수정 정규근사

[예제 6]에서 구하고자 하는 확률은 그림 (a)이며, 정규근사에 의한 근사확률은 그림 (b)와 같다. 따라서 오차를 줄이기 위해 그림 (c)와 같이 수정한 정규근사를 구한다.

$$P(a \le X \le b) \approx P\left(\frac{a - 0.5 - np}{\sqrt{npq}} \le Z \le \frac{b + 0.5 - np}{\sqrt{npq}} \right)$$

$X \sim B(n, p)$

$X \approx N(np, npq)$



(a) (b) (c)

[예제 6] n p

$X \sim B(30, 0.2)$일 때, 확률질량함수를 이용한 $P(X=4)$와 연속성을 수정한 근사확률 $P(X=4)$를 구하라. (p.m.f.)

풀이

$X \sim B(30, 0.2)$

확률질량함수: $f(x) = \binom{30}{x}(0.2)^x(0.8)^{30-x}, \quad x = 0, 1, 2, \cdots, 30$

$\binom{n}{x} p^x (1-p)^{n-x}$

$$P(X=4) = f(4) = \binom{30}{4}(0.2)^4(0.8)^{26} = 0.1325$$

$\boxed{X \approx N(\underset{6}{np},\ \underset{4.8}{npq})}$

$\mu = np = 6$이고 $\sigma^2 = npq = 4.8$이므로 근사적으로 $X \approx N(6, 4.8)$이다. 따라서 3.5와 4.5를 표준화한다. :

$$z_l = (3.5-6)/\sqrt{4.8} = -1.14, \quad z_u = (4.5-6)/\sqrt{4.8} = -0.68$$

$P(X=4)$

$P(3.5 \le X \le 4.5)$

그러면 근사확률은 다음과 같다.

$$\begin{aligned} P(X=4) &= P(3.5 \le X \le 4.5) \approx P(-1.14 \le Z \le -0.68) \\ &= P(0.68 \le Z \le 1.14) = P(Z \le 1.15) - P(Z < 0.68) \\ &= 0.8729 - 0.7517 = 0.1212 \end{aligned}$$