통계 공식과 개념들 한번에 총정리 해드립니다. (이산확률분포, 이항분포, 연속확률분포, 확률밀도함수, 표준정규분포, 표준화공식, 임의추출, 표본평균...

1. 이산확률분포 / 이산확률분포의 평균, 분산, 표준편차

표

→ 떨어져있는 (셀수있는). 확률변수 X.

$E(X)$ $V(X)$ $\sigma(X)$

ex) 주사위던짐

눈. 확률변수 $X = 1, 2, 3, 4, 5, 6$

0.5 0.6 ...

표..

X	1	2	3	4	5	6	합
P(X)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

(X² : 1, 4, 9, 16, 25, 36)

$E(X) = \Sigma X \cdot P(X)$

$= 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + 6\cdot\frac{1}{6}$

$= \frac{1+2+3+4+5+6}{6} = \frac{21}{6}$

$= \frac{7}{2} = 3.5$

$V(X) = E(X^2) - (E(X))^2$

$= \left(1\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 9\cdot\frac{1}{6} + 16\cdot\frac{1}{6} + 25\cdot\frac{1}{6} + 36\cdot\frac{1}{6}\right)$

$- \left(\frac{7}{2}\right)^2$

$= V$

통계 공식과 개념들 한번에 총정리 해드립니다. (이산확률분포, 이항분포, 연속확률분포, 확률밀도함수, 표준정규분포, 표준화공식, 임의추출, 표본평균...

3 (9)	4 (16)	5 (25)	6 (36)	합
$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

$$= \frac{7}{2} = 3.5$$

$$V(X) = E(X^2) - (E(X))^2$$

$$= \left(1\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 9\cdot\frac{1}{6} + 16\cdot\frac{1}{6} + 25\cdot\frac{1}{6} + 36\cdot\frac{1}{6}\right) - \left(\frac{7}{2}\right)^2$$

$$= V$$

$$\sigma(X) = \sqrt{V}$$

(분산에 √만 씌우면 됨).

2. 이항분포 / 이항분포의 평균, 분산, 표준편차

이산확률분포의 일종.

↳ 어떤일 1번 시행할 때 일어날 확률 p (되는확률), 안 되는 확률 q. (매 시행마다 일어날 확률 일정 ex) 동전, 주사위)

n번 시행했을 때.
일어난 횟수를 확률변수 X라 한다.

⇒ 이항분포. $B(n, p)$

ex). ~~주사위~~ 동전 2개 던진다. $p=\frac{1}{4}$, $q=\frac{3}{4}$
5번 던진다. $n=5$.
둘다 앞면이 나오는 횟수 확률변수 X라 하자.
$X=0,1,2,3,4,5$ → $B\left(5, \frac{1}{4}\right)$

$E(X) = np = \frac{5}{4}$

☆ $V(X) = npq = 5\cdot\frac{1}{4}\cdot\frac{3}{4} = \frac{15}{16}$

☆ $\sigma(X) = \sqrt{npq} = \frac{\sqrt{15}}{4}$

$$P(X=2) = {}_5C_2\cdot\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^3$$

$$P(X=r) = {}_nC_r\cdot p^r\cdot q^{n-r}$$

$$\boxed{\begin{array}{ccc} E(X) & V(X) & \sigma(X) \\ \shortparallel & \shortparallel & \shortparallel \\ 5 & 10 & \sqrt{10} \end{array}}$$

$$E(aX+b) = a \cdot E(x) + b.$$

$$V(aX+b) = a^2 V(x)$$

$$E(2X+1) = 2E(x)+1$$

$$= 10+1 = 11$$

$$\sigma(3X-1) = \sqrt{V(3X-1)}$$

$$= \sqrt{9V(x)}$$

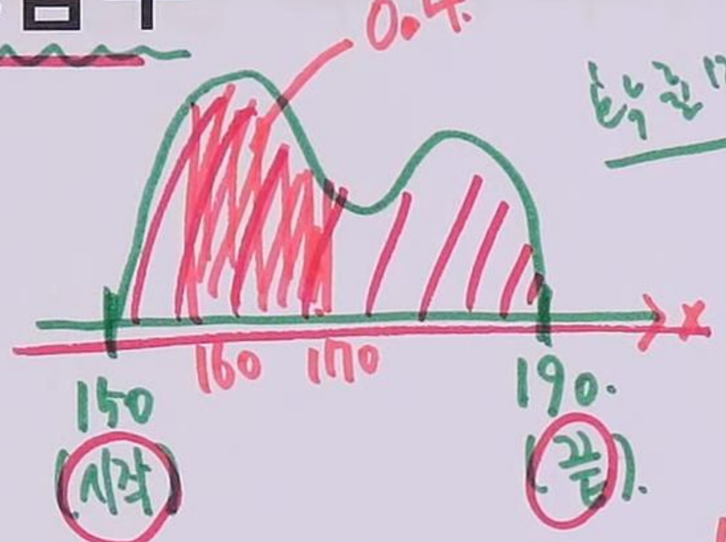
$$= \sqrt{90}.$$

3. 연속확률분포와 확률밀도함수

↳ 확률변수 X가 연속해있다.

ex) 시간. 몸무게. 키. 300명.

함수로!



확률밀도 함수.
(여러모양~)

$P(160 \le X \le 170)$?
= 넓이. = 0.4
= 40%

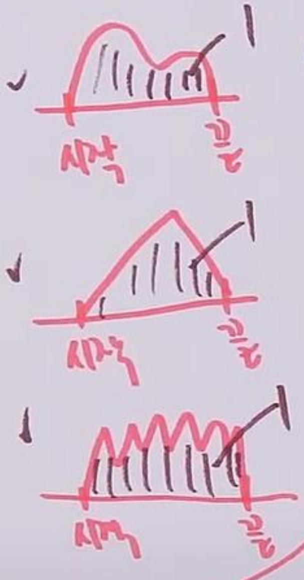
i) 확률은 넓이로~.

ii) 시작~끝까지 확률은 반드시 1~.

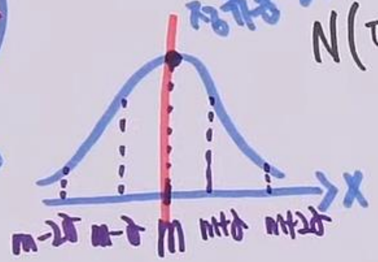
통계 공식과 개념들 한번에 총정리 해드립니다. (이산확률분포, 이항분포, 연속확률분포, 확률밀도함수, 표준정규분포, 표준화공식, 임의추출, 표본평균...

4. 정규분포 / 표준정규분포를 이용한 확률구하기 / 표준화공식

연속확률분포 정의



정규분포.

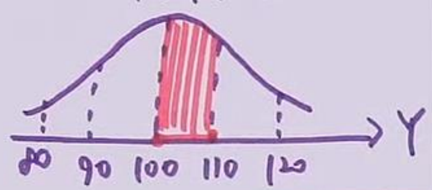


$N(\underset{m}{\text{평균}}, \underset{\sigma}{\text{표준편차}}^2)$

$N(m, \sigma^2)$

변수X, $N(10, 2^2)$

변수Y, $N(100, 10^2)$



변수Z, $N(0, 1^2)$

표준정규분포.

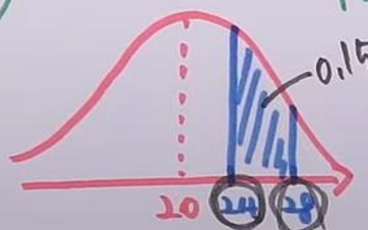
$(P(0 \le Z \le 1) = 0.3413)$

$P(10 \le X \le 12) = 0.3413$

$= P(0 \le Z \le 1)$

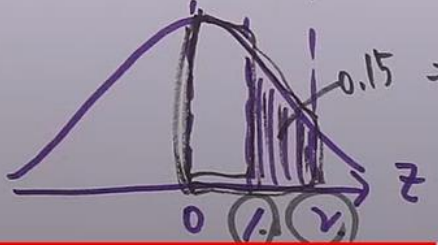
ex). 변수X인 $N(20, 4^2)$이 있다.

이때, $P(24 \le X \le 28)$?



$\left(\begin{array}{l} P(0 \le Z \le 1) = 0.3 \\ P(0 \le Z \le 2) = 0.45 \end{array}\right)$

$P(0 \le Z \le 2) - P(0 \le Z \le 1)$



$0.15 = 0.45 - 0.3 = 0.15$

$\frac{24-20}{4} = \frac{4}{4} = 1$

$\frac{28-20}{4} = \frac{8}{4} = 2$

표준화공식

$\frac{X - m}{\sigma}$

5. 이항분포 to 정규분포

$$B(n, p) \xrightarrow{n \geq 30 \text{이면}} N(m, \sigma^2)$$

ex) $B\left(100, \frac{1}{4}\right) \longrightarrow N\left(25, \left(\frac{5\sqrt{3}}{2}\right)^2\right)$

$E(X) = 100 \cdot \frac{1}{4} = 25$

$V(X) = 100 \cdot \frac{1}{4} \cdot \frac{3}{4} = 25 \cdot \frac{3}{4} = \frac{75}{4}$

$\sigma(X) = \sqrt{\frac{75}{4}} = \frac{5\sqrt{3}}{2}$

6. 임의추출 n 개를 하면?

변수 X $\xrightarrow[\text{(표본을 뽑는다)}]{\text{임의추출 } n\text{개.}}$ 변수 $\overline{X}$ (표본평균)

$E(X) = m$	그대로.	$E(\overline{X}) = m$
$V(X) = \sigma^2$	n으로 나눈다.	$V(\overline{X}) = \dfrac{\sigma^2}{n}$
$\sigma(X) = \sigma$	$\sqrt{n}$ 으로 나눈다.	$\sigma(\overline{X}) = \dfrac{\sigma}{\sqrt{n}}$.

7. 통계적 추정 / 모평균의 추정(추측).

m

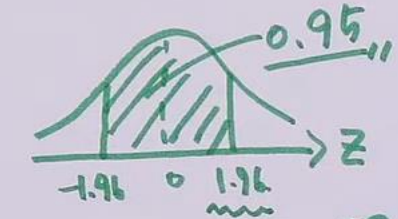
n명 임의추출 ~

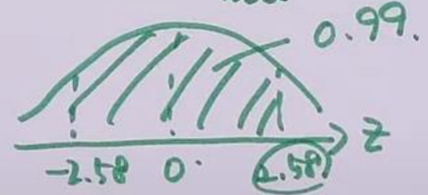
↳ 표본평균 $\bar{X}$

$$\bar{X} - k\cdot\frac{\sigma}{\sqrt{n}} \le m \le \bar{X} + k\cdot\frac{\sigma}{\sqrt{n}}$$ ← 신뢰구간.

σ : 모 표준편차 ($n \ge 30$ 이면, $\sigma = \bar{\sigma}$)

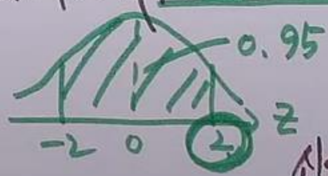
k : 신뢰도 상수.
- 95% ⇒ k=1.96.
- 99% ⇒ k=2.58





☆ 신뢰도상수는 반드시 문제에서 주어짐.

ex). 100명 임의추출 했더니 $\bar{X} = 20$, $\bar{\sigma} = 3$, 신뢰도 95%로 m을 추정한 신뢰구간을 구하시오. ($P(-2 \le Z \le 2) = 0.95$)



$$20 - 2\cdot\frac{3}{\sqrt{100}} \le m \le 20 + 2\cdot\frac{3}{\sqrt{100}}$$

$$20 - \frac{6}{10} \le m \le 20 + \frac{6}{10}$$

$$\frac{196}{10} \le m \le \frac{206}{10} \rightarrow 19.6 \le m \le 20.6$$

신뢰구간.

신뢰구간 길이
$20.6 - 19.6 = 1$