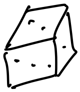
Ch 01. 정규분포

1. 확률의 기초
2. 정규 분포
3. 중심 극한 정리
4. 표본 합의 확률 분포
5. 우리 주변의 정규분포

1.1 확률의 기초

1.1.1 확률 변수와 확률 분포

확률 변수 $= x$

 → $x = \{1, 2, 3, 4, 5, 6\}$ → 표본 or 샘플

$\{\frac{1}{6}, \dots p(x=3), \dots, \frac{1}{6}\}$ → 확률 분포

모두 같은 확률 → 균등 분포

확률 분포가 되기 위한 조건

이산 확률 분포

1. $0 \le p(x_k) \le 1 \quad (k=1, \dots, N)$
2. $\sum_{k=1}^{N} p(x_k) = 1$

연속 확률 분포

1. $p(x) \ge 0$
2. $\int_{-\infty}^{\infty} p(x)\,dx = 1.$

1.1.2 확률 분포의 종류

이산 확률 분포 ⇒ x가 이산값 ⇒ 

연속 확률 분포 ⇒ x가 연속값 ⇒ 

1.1.3 기댓값과 분산

기댓값 = 평균

이산 $E[x] = \sum_{x=1}^{N} x_k p(x_k)$

연속 $E[x] = \int_{-\infty}^{\infty} x p(x)\,dx$

분산 = $(x-\mu)^2$의 기댓값

$$Var[x] = E[(x-\mu)^2] = \sum_{k=1}^{N} (x_k-\mu)^2 p(x_k)$$

$$Var[x] = E[(x-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)\,dx$$

1.2 정규 분포

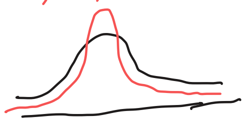
정규 분포 $= p(x;\mu,\sigma) = N(x;\mu,\sigma) = \dfrac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\dfrac{(x-\mu)^2}{2\sigma^2}\right)$

확률 변수 매개변수

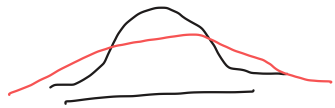
만약 $\mu=0,\ \sigma=1$ 로. $N(x;0,1) = \dfrac{1}{\sqrt{2\pi}} \exp\left(-\dfrac{x^2}{2}\right)$ ⇒ "표준 정규 분포"

if) $\mu \Rightarrow \uparrow$

if) $\sigma \Rightarrow \downarrow$



if) $\sigma \Rightarrow \uparrow$



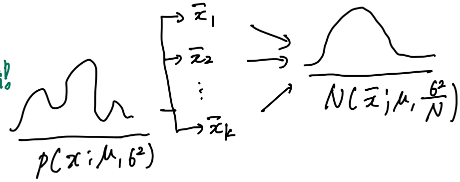
1.3 중심 극한 정리

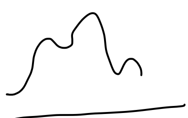
모든 확률 분포 $p(x) \xrightarrow{\text{추출!}} \{x^{(1)}, x^{(2)}, \cdots, x^{(N)}\}$

일수와방

표본 평균 $= \bar{x} = \dfrac{x^{(1)} + x^{(2)} + \cdots + x^{(N)}}{N}$

즉!





$p(x;\mu,\sigma^2)$

⇒

if) N=1

$N(\bar{x};\mu,\frac{\sigma^2}{1})$

if) N=2

$N(\bar{x};\mu,\frac{\sigma^2}{4})$

. . . if) N=100↑ 더 커지면?

$N(\bar{x};\mu,\frac{\sigma^2}{N})$

1.4 표본합의 확률분포

$$\bar{x} = \frac{x^{(1)} + x^{(2)} + \cdots + x^{(N)}}{N} \xrightarrow{\text{표본합}} S = x^{(1)} + x^{(2)} + \cdots + x^{(N)} = N\bar{x}$$

즉! ① $E[S] = E[N\bar{x}] = NE[\bar{x}] = \boxed{N\mu}$ ☆

② $Var[S] = Var[N\bar{x}] = E[(N\bar{x} - N\mu)^2] = \boxed{N\sigma^2}$

규칙 발견!

$p(x; \mu, \sigma^2)$ 일 때!

이러면 분산이 작아져!

$p(\bar{x}; \mu, \frac{\sigma^2}{N})$ 이고!

근데! 이러면 평균과 분산 모두 ↑

$p(S; N\mu, N\sigma^2)$ 이다!