3.6 Chevyshev Inequality

mean = ? $E[X]$

- estimate / prediction : $\hat{X}$ → $(X-\hat{X})^2$ → $E[(X-\hat{X})^2]$ → minimize expectation of squared difference.
- (real) random observation : X →

↳ 이 값은 작을 수록 유리하다. ⇒ $\hat{X}$? $= E[X]$

X는 random 하고 뭐가 나올지 모르니까 한 번만 해서는 $\hat{X}$이 좋은 줄 모른다.

그래서 여러번 해서 평균 값을 구해준다.

$$f(\hat{x}) = \frac{(\hat{x}-x_1)^2 + (\hat{x}-x_2)^2 + \cdots + (\hat{x}-x_n)^2}{n} \qquad \hat{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

$$P(|X - \underbrace{E[X]}_{\hat{X}}| \geq a) \leq \frac{\sigma_x^2}{a^2}$$

ex) $a = 2\sigma_x$

$$P(|X-E[X]| \geq 2\sigma_x) \leq \frac{1}{4}$$

유도,

↱ 이것의 가장 minimum 값은 a,,

↱ 이것이 a 보다 작다는 설정이니까 0이라고 생각하자.

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - E[X])^2 f_x(x)\,dx = \int_{|X-E[X]| \geq a} (x - E[X])^2 f_x(x)\,dx + \int_{|X-E[X]| < a} (x - E[X])^2 f_x(x)\,dx$$

$$\geq \int_{|X-E[X]| \geq a} a^2 f_x(x)\,dx$$

$$\Downarrow$$

$$\sigma_x^2 \geq \int_{|X-E[X]| \geq a} a^2 f_x(x)\,dx \Rightarrow \frac{\sigma_x^2}{a^2} \geq \int_{|X-E[X]| \geq a} f_x(x)\,dx$$

density를 어떤 구간에 대해서 적분하면

그 조건을 만족하는 영역에 속할 확률.

↳ $P(|X-E[X]| \geq a) \leq \frac{\sigma_x^2}{a^2}$

Chap 4. Special Dist.

4.2 Bernoulli Dist.

RV : $X \rightarrow$ binary

$$\begin{cases} P(\text{success}) = p \longrightarrow 1 \\ P(\text{failure}) = 1-p \longrightarrow 0 \end{cases} \qquad E[X] = p,\ \sigma_X^2 = p(1-p)$$

4.3 Binomial Dist (이항분포)

$\Rightarrow$ X : # of success out of n Bernoulli trial

$x = 0, 1, 2, \cdots, n$ (discrete)

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

1) $\sum_{x=0}^{n} P_X(x) = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} = (p + (1-p))^n = 1.$

이항정리, $(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x}$

$$E[X] = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} \frac{x \cdot n!}{(n-x)!\,x!} p^x (1-p)^{n-x} = \sum_{x=1}^{n} \frac{n!}{\underbrace{(n-x)!}_{(n-1-(x-1))}\,(x-1)!} p^x (1-p)^{n-x} = \sum_{x=1}^{n} \frac{n \cdot (n-1)!\, p}{(n-1-(x-1))!\,(x-1)!} \cdot p^{x-1} (1-p)^{(n-1)-(x-1)}$$

$x-1$ 을 x' 으로 치환.

$$(a+b)^{n-1} = \sum_{x=0}^{n-1} \frac{(n-1)!}{(n-1-x)!\,x!} p^x (1-p)^{n-1-x}$$

$$\Rightarrow \sum_{x'=0}^{n-1} np \cdot \frac{(n-1)!\, p^{x'} (1-p)^{n-1-x'}}{(n-1-x')!\,x'!}$$

모양이 똑같다.

$$= np\,(p + (1-p))^{n-1}$$

$$= np.$$

$$E[X^2] = \sum_{x=0}^{n} \frac{x^2\, n!}{(n-x)!\,x!} p^x (1-p)^{n-x} = \sum_{x=1}^{n} \frac{\overset{(x-1+1)}{x}\, n!}{(n-x)!\,(x-1)!} p^x (1-p)^{n-x} = \sum_{x=1}^{n} \frac{\cancel{(x-1)}\, n!}{(n-x)!\,\underbrace{\cancel{(x-1)!}}_{(x-2)!}} p^x (1-p)^{n-x} + \underbrace{\sum_{x=1}^{n} \frac{n!}{(n-x)!\,(x-1)!} p^x (1-p)^{n-x}}_{np}$$

$$= \sum_{x=2}^{n} \frac{n!}{(n-x)!\,(x-2)!} p^x (1-p)^{n-x} + np$$

$$= \sum_{x=2}^{n} \frac{n(n-1)(n-2)!}{((n-2)-(x-2))!\,(x-2)!} p^2 p^{x-2} (1-p)^{(n-2)-(x-2)} + np$$

$$x-2 \rightarrow x' \quad \Rightarrow \underbrace{\sum_{x'=0}^{n-2} \binom{n-2}{x'} p^{x'} (1-p)^{n-2-x'}}_{} \cdot n(n-1)p^2 + np$$

$$(p + (1-p))^{n-2}\, n(n-1)p^2 + np$$

$$\sigma_X^2 = E[X^2] = n^2p^2 = np(1-p)$$

$$(pt + (1-p))^n = \sum_{x=0}^{n} \binom{n}{x} p^x t^x (1-p)^{n-x}$$

$\frac{d}{dt} \Rightarrow np(pt + (1-p))^{n-1} = \sum_{x=0}^{n} x t^{x-1} \binom{n}{x} p^x (1-p)^{n-x}$ 이항정리에서 미분을 이용하는 것은

$t=1$ $np(p+(1-p))^{n-1} = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$ 평균을 구할 때 쓸 수 있다.

$$np = E[X]$$

$\frac{d^2}{dt^2} \Rightarrow n(n-1)p^2(pt + (1-p))^{n-2} = \sum_{x=0}^{n} x(x-1)t^{x-2} \binom{n}{x} p^x (1-p)^{n-x}$

$t=1$ $(n^2-n)p^2 = \sum_{x=0}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x} - \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$

$$\Rightarrow E[X^2] - E[X]$$

4.4 Geometric Dist

X : # of Bernoulli trials until 1st success.

$P_X(x) = (1-p)^{x-1} \cdot p \ , \ x = 1, 2, 3, \cdots$

$E[X] = \frac{1}{p} \ , \ \sigma_X^2 = \frac{1-p}{p^2}$

◎ Forgetfulness (Memoryless)

주사위가 '6'이 나올 때까지 계속 던진다. '6'이 나오면 끝나는 것. ⇒ 가장 대표적인 Geometry Dist.

10번 던졌는데 '6'이 안나왔다. 앞으로 5번안에 '6'이 나올 확률은? 앞에서 5번을 실패했든, 10번을 실패했든 그 숫자가 전혀 중요하지 않다.

5번 던졌는데 '6'이 안나왔다. 앞으로 5번안에 '6'이 나올 확률은? → 잊어버린다. → Forgetfulness.

⇒ Consider K additional trials until the 1st success, given n trials fail.

현재 n번까지 try를 했는데 전부 실패했다. K번 더 try를 해서 K번째에 성공하게 될 확률은?

$$P(X = n+K \mid X > n) = \frac{P(X = n+K \cap X > n)}{P(X > n)} = \frac{P(X = n+K)}{P(X > n)} = \frac{(1-p)^{n+K-1} p}{\sum_{x=n+1}^{\infty} (1-p)^{x-1} \cdot p} = \frac{(1-p)^{n+K-1} p}{\frac{p(1-p)^n}{1-(1-p)}} = p(1-p)^{K-1}$$

$$= P(X = K)$$