$$E[X] = \begin{cases} \sum x_i P(x_i) \\ \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx \end{cases} \qquad w_i \rightarrow 0 < w_i < 1, \quad \sum w_i = 1$$

$$E[(X-\mu)^2] = E[X^2] - \mu^2$$

Ex 3.13 > Geometric distribution

Random Variable K : # of trials until the 1st success.

첫 번째로 성공할 때까지 시행한 횟수.

ex) tossing a die repeatedly until 1st "6"

$K=1 \rightarrow p$

$K=2 \rightarrow (1-p)p$

$K=3 \rightarrow (1-p)(1-p)p$

$\vdots$

$$P_K(k) = (1-p)^{k-1} \cdot p, \quad k = 1, 2, 3, 4, \cdots$$

1부터 시작.

(a) $\sum_{k=1}^{\infty} p(1-p)^{k-1} = 1$

(b) $P(K>5) = \sum_{k=6}^{\infty} p(1-p)^{k-1}$

$P(K \le 5) = F_K(5) = \sum_{k=1}^{5} p(1-p)^{k-1}$ ↝ 확률값만 정확하게 구하면, discrete case를 정확하게 규정할 수 있다.

(c) $E[K] = \sum_{k=1}^{\infty} k\,p(1-p)^{k-1} = p \cdot \frac{1}{p^2} = \frac{1}{p}$

$\sum_{k=1}^{\infty} (1-p)^k = \frac{1-p}{p}$ (무한 등비 급수)

$\downarrow \frac{d}{dp}$

$+\sum_{k=1}^{\infty} k(1-p)^{k-1} = +\frac{1}{p^2}$

$\xrightarrow{\frac{d}{dp}} -\sum k(k-1)(1-p)^{k-2} = -\frac{2}{p^3}$

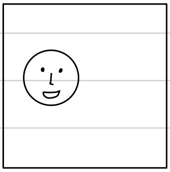
$\Rightarrow \sum_{k=1}^{\infty} k^2(1-p)^{k-2} = \sum k(1-p)^{k-2} + \frac{2}{p^3}$

$\times\, p(1-p) \qquad \times\, p(1-p)$

$\sum_{k=1}^{\infty} k^2(1-p)^{k-1} p = \sum_{k=1}^{\infty} k(1-p)^{k-1} p + \frac{2(1-p)}{p^2}$

$\sigma_K^2 = E[K^2] - \mu^2 = \frac{1-p}{p^2}$

$\sum_{k=1}^{\infty} k^2(1-p)^{k-1} \cdot p = \frac{1}{p} + \frac{2(1-p)}{p^2}$



Digital camera 에 얼굴을 찾아내는 기능이 있다.

얼굴인지 아닌지 판단하려고 할 때, 무엇을 가지고 판단할까?

$\sum$(표준 참값 − 실제값)^2 → minimum.

참값 : 예측, 추정

↳ error model

(값)^2 → energy.

" 차이에 해당하는 energy가 minimum 이 되도록 하자! "

⇒ 무엇인가를 판단하고 선택하는 기준은 직관적으로 하는 것이 아니라, 수치 모델을 세워서 수치 모델에 해당하는 값이 가능하면 일정한 기준보다 작을 수록 '좋다'라고 판단하는 것.

random 하게 나오는 값.

$$\min_x f(x) = P(a_1)(x-a_1)^2 + P(a_2)(x-a_2)^2 + P(a_3)(x-a_3)^2 + \cdots + P(a_n)(x-a_n)^2$$

⇒ error model 이자, energy.

예측하려는 값.

ㄴ눈이 잘 맞는 것. ㄴ코가 잘 맞는 것. ㄴ입이 잘 맞는 것.

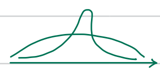
f(x)를 최소로 만드는 x 값 : a_1 , $\frac{a_1+a_2}{2}$ . . . $\hat{x} = \frac{\sum_{i=1}^{n} a_i}{n}$ 일 때 f(x)가 최소가 된다.

$P(a_i)$

≀ 평균!

$$f(\hat{x}) = \frac{\sum_{i=1}^{n}(a_i-\hat{x})^2}{n} \rightsquigarrow n\sigma_A^2 \rightsquigarrow \text{분산!}$$

Data 가 얼마나 퍼져있는가. 정확도를 예측 하려면, 분산이 작은게 편하다.



energy model 이 minimum이 되도록 만들어주는 것 : data 값들의 평균.

평균의 값들을 대입했을 때 minimum energy는 분산이 된다.

분산이 작을 수록 예측하는 값의 범위가 비교적 일정하다. (좁은 범위). 그래서 더 정확해 진다.

눈에 해당하는 위치의 error 값.
코에 해당하는 위치의 error 값.
입에 해당하는 위치의 error 값.

눈은 잘 맞는데, 입이 잘 안맞을 수도 있다.
동시에 잘 맞는 x 값을 찾는 것이다.
어느 하나는 잘 안맞을지 모르지만, 전체적으로 잘 정렬된 얼굴 위치를 찾아준다.

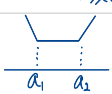
판단, 선택해야 하는
⇒ 문제는 이런 model을 쓴다.

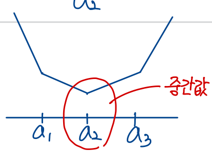
Median : 데이터 값들의 중간 값. 1, 2, $\underset{\text{median}}{4}$, 5, 7

$$\min g(x) = P(a_1)|x-a_1| + P(a_2)|x-a_2| + P(a_3)|x-a_3|$$

↝ '확률' 같은 가중치가 붙으면 단순하게 a_1, a_2, a_3 의 위치만으로 판단할 수 없고, 가중치를 따져서 0.5 정도의 확률 위치에 나오는 중간값이 median 이 된다.

g(x)를 최소로 만드는 x값 : a_1 a_1과 a_2 사이에 있는 모든 값. a_2





3.5 Conditional Mean

$$E[X|A] = \sum_{x_i \in A} x_i P(x_i|A) = \sum_{x_i \in A} x_i \frac{P(x_i \cap A) \to P(x_i)}{P(A)}$$ ⇝ discrete 한 경우 Conditional mean 이라고 이야기 한다.

새로운 sample space A를 벗어나지 않는 변수값들에 대해서만

조건부 확률을 구해서 평균을 구한다.

for continuous case : conditional mean 이 복잡해진다.

$$E[X|A] = \int_{x \in A} x \cdot f_X(x|A)\,dx$$

conditional density를 구해야 한다.

이것이 문제.

$$f_X(x|A) = \frac{d}{dx} F_X(x|A)$$ ⇝ Conditional density는 Conditional CDF 에서 구한다. CDF를 미분하면 density가 나오므로.

$$= \frac{d}{dx} P(X \le x \mid A)$$ → Conditional CDF를 확률로 바꾼 것.

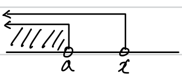
$$= \frac{d}{dx} \frac{P(X \le x \cap A)}{P(A)}$$

" A 라는 event 가 어떻게 주어지느냐에 달려 있다. "

Ex 3.14 > $A = \{X \le a\}$

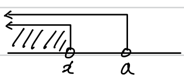
$$f_X(x|x \le a) = \frac{d}{dx} \frac{P(X \le x \cap X \le a)}{P(X \le a)}$$

(i) $x > a$



$$\frac{d}{dx} \frac{P(X \le x \cap X \le a)}{P(X \le a)} \Rightarrow \frac{d}{dx} \frac{P(X \le a)}{P(X \le a)} = \frac{d}{dx} \cdot 1 = 0$$

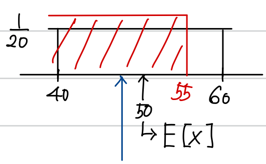
(ii) $x \le a$



$$\frac{d}{dx} \frac{P(X \le x \cap X \le a)}{P(X \le a)} \Rightarrow \frac{d}{dx} \frac{P(X \le x)}{P(X \le a)} \Rightarrow \frac{d}{dx} \frac{F_X(x)}{F_X(a)} = \frac{f_X(x)}{F_X(a)}$$

Ex 3.15 > $f_X(x) = \frac{1}{20}$, $40 \le x \le 60$ $A = \{X \le 55\}$

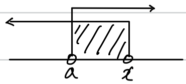
$$E[X|A] = \int_{x \in A} x \cdot f_X(x|A)\,dx = \int_{40}^{55} x \frac{\frac{1}{20}}{F_X(55)}\,dx = 47.5$$



47.5 conditional mean 의 값.

$A = \{ X > a \}$

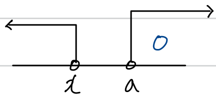
$$F_X(x \mid A) = \frac{P(X \le x \cap X > a)}{P(X > a)}$$



$$\Rightarrow \frac{P(a < X < x)}{1 - F_X(a)}$$

$$= \frac{F_X(x) - F_X(a)}{1 - F_X(a)}$$

$A = \{ x \le a \}$



$$\text{CDF} = \frac{0}{P(X > a)} \qquad f_X(x \mid A) = \frac{f_X(x)}{1 - F_X(a)}$$