CHAP.3
확률론, 확률변수, 기댓값

nonezero@kumoh.ac.kr

확률종류

- 고전적(수학적) 확률
- 통계적(경험적) 확률

$$P(A) = \frac{|A|}{|S|}$$

$$P(A) = \frac{\text{사건 } A\text{의 도수}}{\text{총 관찰 도수}}$$

시행 횟수에 따라 달라진다.

경우의 수

- 3장의 카드 중에서 3장을 뽑아서 **순서**대로 **나열**하는 경우의 수

  ABC, ACB, BAC, BCA, CAB, CBA　　3가지, 2가지, 1가지　　$3! = 3 \times 2 \times 1,$　　$0! = 1$

- 3장의 카드 중에서 2장을 뽑아서 **나열**하는 경우의 수

순서 ← → 무순

AB　BA
AC　CA
BC　CB

두 개를 나열하는
경우의 수

$$\mathbf{3 \times 2}$$

$$nPr = \frac{n!}{(n-r)!}$$

Permutation
순열

$$\frac{\mathbf{3 \times 2}}{\mathbf{2 \times 1}}$$

$$nCr = \frac{nPr}{r!} = \frac{n!}{(n-r)!\,r!}$$

Combination
(조합)

확률

기본정의

- 확률실험
- 근원사건
- 표본공간
- 사건

$$S = \{\text{모든 근원사건}\}$$

$$A = \{\text{일부 근원사건}\}$$

집합

$$A \subset S$$

주사위 1개 던지는 확률실험

{1} {2} {3} {4} {5} {6}

{1, 2, 3, 4, 5, 6}

짝수가 나오는 사건
{2, 4, 6}

확률 (Probability)

- 어떤 사건이 발생할 가능성의 '크기'
- 0에서 1사이

확률공식

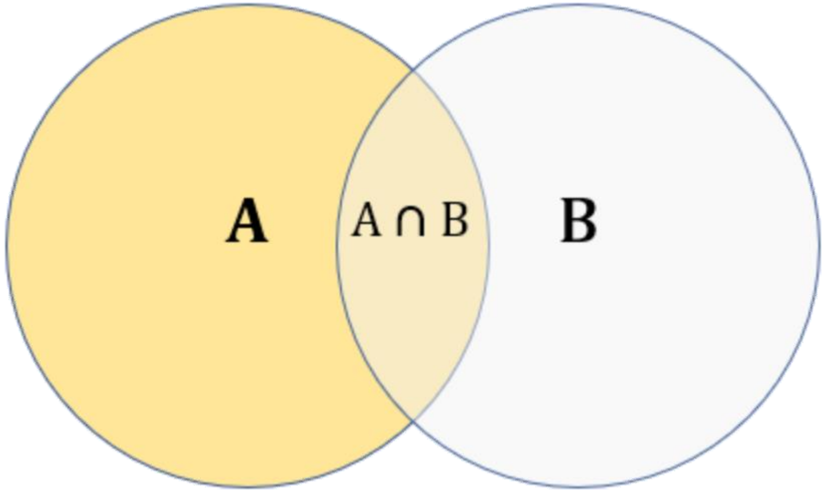
- 집합 연산
  - AND OR NOT (교집합, 합집합, 여집합)

- P(S) = 1
- P(A OR B) = P(A) + P(B) – P(A AND B)
- P(A) = 1 – P(NOT A)

- P(A|B) = ?   Probability of A given B

조건부 확률

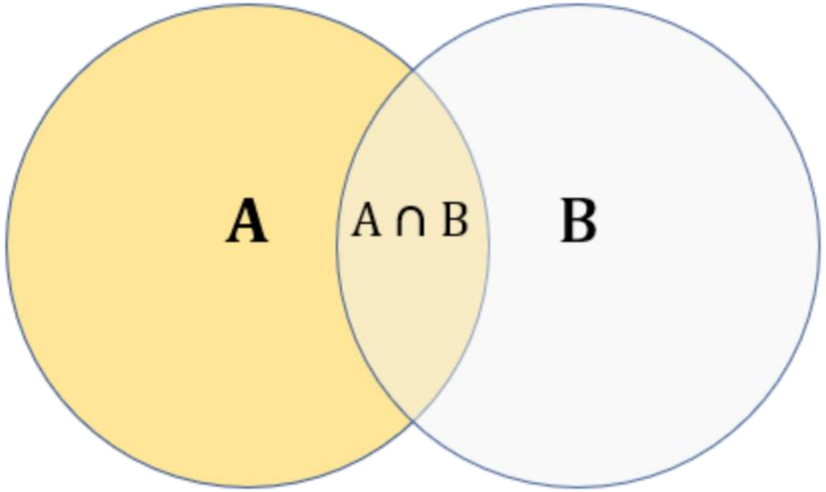
- 사건 A가 발생했을 때, 사건 B가 발생할 확률



$$\frac{P(B \cap A)}{P(A)}$$

조건부 확률 (Conditional Probability)

- 사건 A가 발생했을 때, 사건 B가 발생할 확률



$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

P of B given A

$$S=\{1, 2, 3, 4, 5, 6)$$

짝수 사건 $A = \{2, 4, 6\}$

3의 배수 사건 $B = \{3, 6\}$

$$A \cap B = \{6\}$$

$$P(A \cap B) = \frac{1}{6}$$

$$P(A) = \frac{3}{6}$$

$$\boxed{P(B|A)} = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}$$

짝수 사건일 때
3의 배수 사건

- 흰색, 검은색 주사위 던지는 실험,

- 사건 A: 두 주사위의 나온 수를 합해서 3
- 사건 B: 흰색 주사위가 1
- 사건 A의 확률
- 사건 B의 확률
- 사건 B가 일어난 조건에서 사건 A가 일어날 확률

$$P(A) = \frac{2}{36}$$

$$P(B) = \frac{1}{6}$$

$$P(A|B) = \frac{1}{6}$$

= 검은색 주사위가 2일 확률

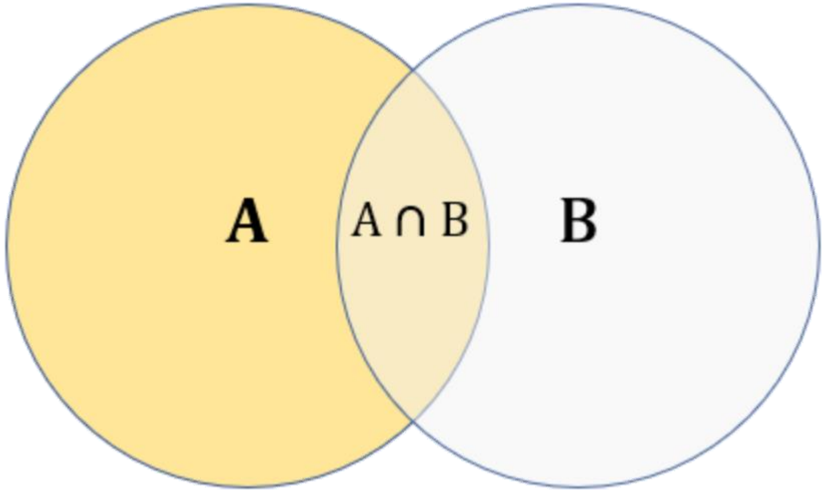
$$P(A|B) = \frac{P(A\ AND\ B)}{P(B)} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6}$$

{(1,1)}{(1,2)}{(1,3)}{(1,4)}{(1,5)}{(1,6)}

{(2,1)}{(2,2)}

교집합에 주목해 보자

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$



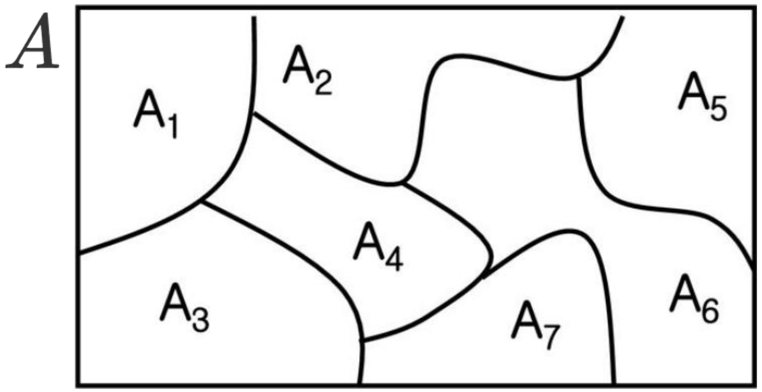
$$P(B \cap A) = P(B|A)P(A)$$

$$P(A \cap B) = P(A|B)P(B)$$

베이즈 정리 (Bayes Rule / Theorem)

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Law of Total Probability

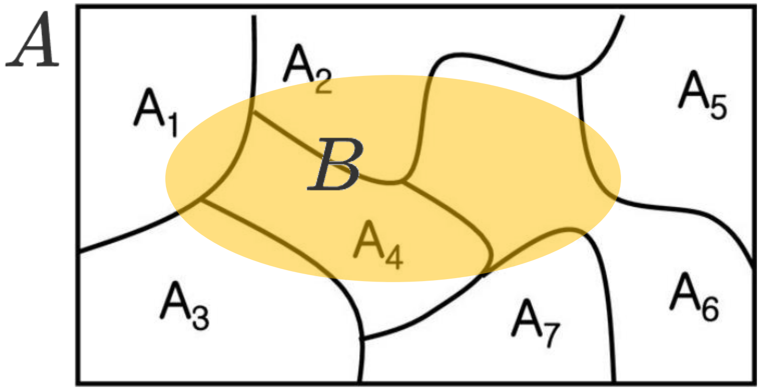


이런 상황이라면,

$$P(A) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

$$P(A_i \cap A_j) = 0 \text{ for } i \neq j$$

$$P(\cup_{i=1}^{n} A_i) = 1$$



B가 이럴 때,

$$\boxed{P(B)} = P(B \cap A_1) + P(B \cap A_2) + \cdots + P(B \cap A_n)$$

$$P(B \cap A) = P(B|A)P(A)$$

$$= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \cdots + P(B|A_n)P(A_n)$$

$$\boxed{= \sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

베이즈 정리 (Bayes Rule / Theorem)

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$= \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

likelihood Prior

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

posterior

Evidence

이럴 때 사용

- 인구 1000명당 한 명이 걸리는 희귀병
- 검사법이 있다; 병에 걸린 사람의 경우 99% 양성반응
- 병에 걸리지 않은 사람의 경우 2% 양성반응
- 홍길동의 검사결과로 양성반응이 나타났을 때,
- 홍길동이 병에 걸렸을 확률은?

양성반응이 나타난 경우 감염되었을 확률

- 사건 A : 피검자가 병에 걸려 있다.
- 사건 B : 피검자가 양성반응을 보인다.

$$P(A|B)$$

주어진 확률 정리

- 인구 1000명당 한 명이 걸리는 희귀병
- 검사법이 있다; 병에 걸린 사람의 경우 99% 양성반응
- 병에 걸리지 않은 사람의 경우 2% 양성반응
- 홍길동의 검사결과로 양성반응이 나타났을 때,
- 홍길동이 병에 걸렸을 확률은?

$$P(A) = 0.001$$
$$P(NOT\ A) = 0.999$$
$$P(B|A) = 0.99$$
$$P(B|NOT\ A) = 0.02$$

풀이

- $P(A) = 0.001$
- $P(NOT\ A) = 0.999$
- $P(B|A) = 0.99$
- $P(B|NOT\ A) = 0.02$
- $P(B) = P(B\ AND\ A) + P(B\ AND\ NOT\ A)$
- $P(B\ AND\ A) = P(B|A)P(A) = 0.99 \times 0.001 = 0.00099$
- $P(B\ AND\ NOT\ A) = P(B|NOT\ A)P(NOT\ A) = 0.02 \times 0.999 = 0.01998$

$$P(A|B) = \frac{P(A\ AND\ B)}{P(B)} = \frac{P(B|A)\ P(A)}{P(B)} = \frac{P(B|A)\ P(A)}{P(B|A)P(A) + P(B|NOT\ A)\ P(NOT\ A)}$$

$= 0.0472$

Bayes' Theorem

- $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

Bayesian Inference ≈ Posterior 구하기

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

양성일 때,
병에 걸렸을 확률

$\propto P(B|A)P(A)$

판정만 할 경우,
$P(B)$를 계산할 필요가 없다.

$P(B|A)P(A) \longleftrightarrow P(B|A')P(A')$

- **Posterior**: $P(A|B)$
- **Prior**: $P(A)$ 병에 걸릴 확률
- **Likelihood**: $P(B|A)$ 병에 걸렸을 때 양성이 나타날 확률
- Evidence: $P(B)$ 양성이 나타날 확률

두 사건이 독립

두 사건이 독립
A 사건이 발생할 확률은 B사건이 발생하든 말든...상관없다.

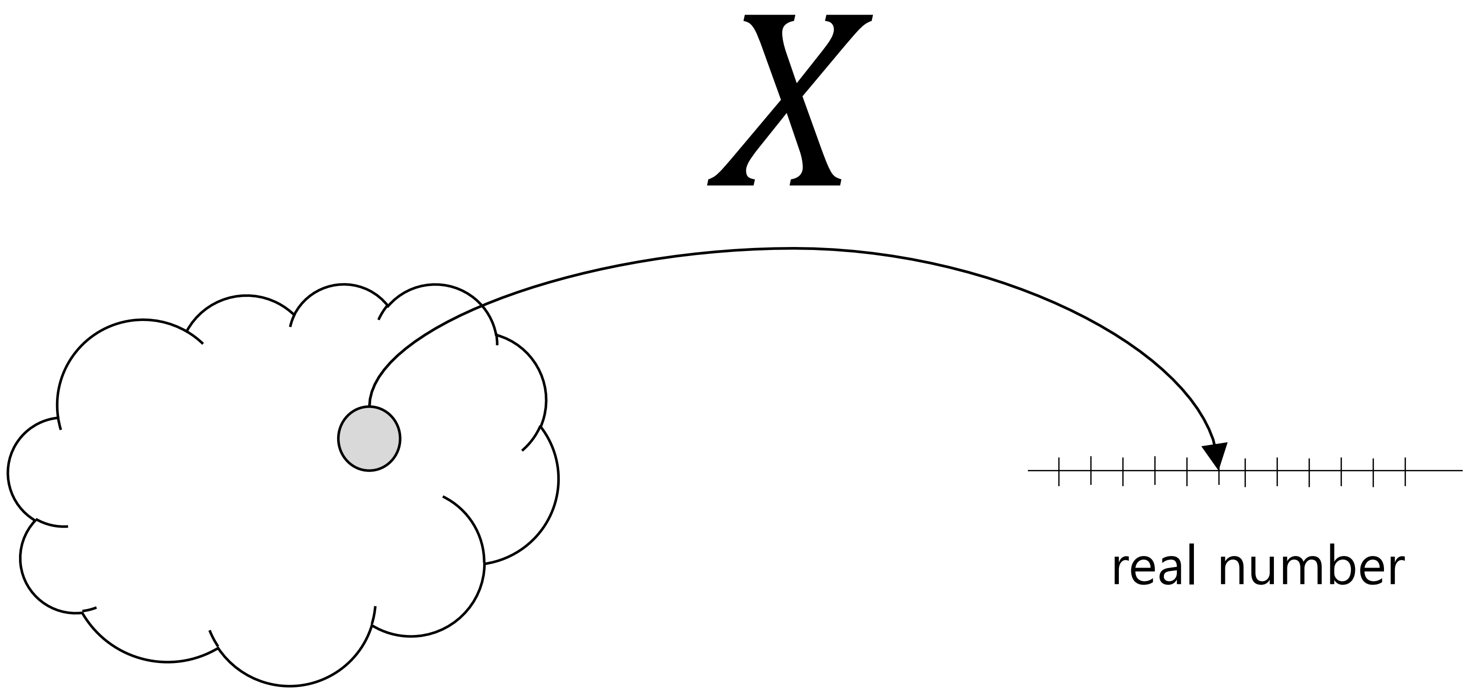
- $P(A|B) = P(A)$

- $P(A\ AND\ B) = P(A|B)P(B) = P(A)P(B)$

확률변수

이제부터는 '사건' 대신
'확률변수'를 쓰자

Random Variable; 확률변수



※ 확률실험, 근원사건, 표본공간, 사건, 실수

Random Variable

- mapping function: event → real value
- $P(X = x)$: X가 x값을 가질 확률
- $P_X(x)$; $P(x)$

- 동전 2개를 던져서 앞면이 나오는 횟수를 확률변수 X라 하자.
- X 가 가질 수 있는 값은? $0, 1, 2$
- $P(X = 1) = P_X(1) = P(1) = {}^1/_2$

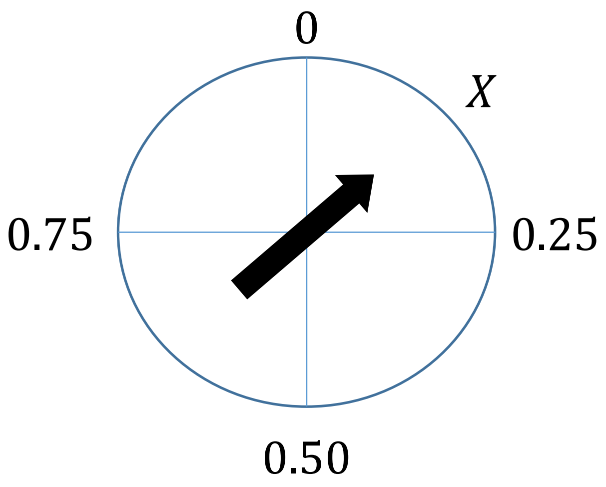
확률변수, 확률을 쫙 펼쳐(distribute)보자

x	0	1	2
$P(X = x)$	$P(X = 0) = {}^{1}/_{4}$	$P(X = 1) = {}^{1}/_{2}$	$P(X = 2) = {}^{1}/_{4}$

확률분포표

연속확률변수

- $P(X = x) = 0 \;\; for\ all\ x$
- X가 취할 수 있는 값이 무한 개

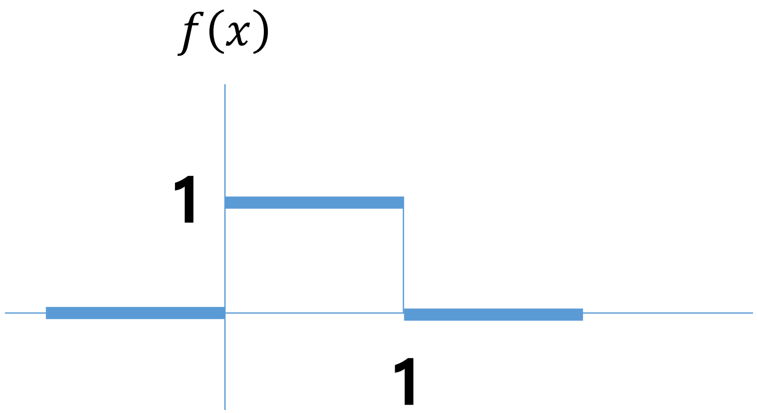


$$P(0.25 \le X \le 0.75) = 0.5$$

$$P(X = 0.5) = {}^{1}\!/_{\infty} = 0$$

그래프; 함수로 그려보자

- 확률분포표를 그릴 수는 없으니까.



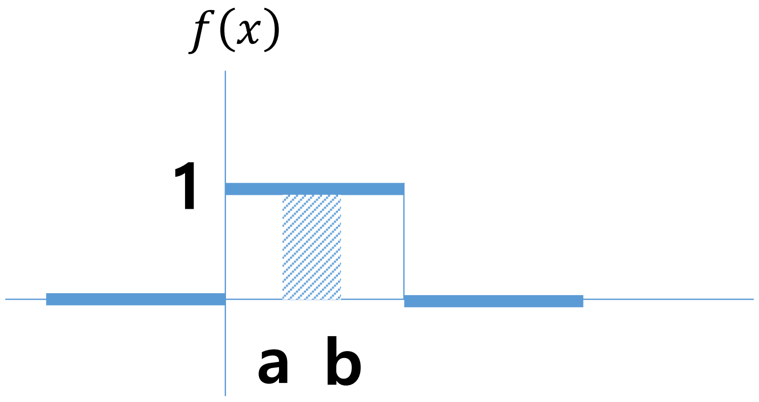
$$f(x) = 0 \qquad when \;\; x < 0$$

$$f(x) = 1 \qquad when \;\; 0 \leq x \leq 1$$

$$f(x) = 0 \qquad when \;\; x > 1$$

$f(x)$라는 함수가 하나 도입된 것에 주목

a, b사이의 면적 = 확률



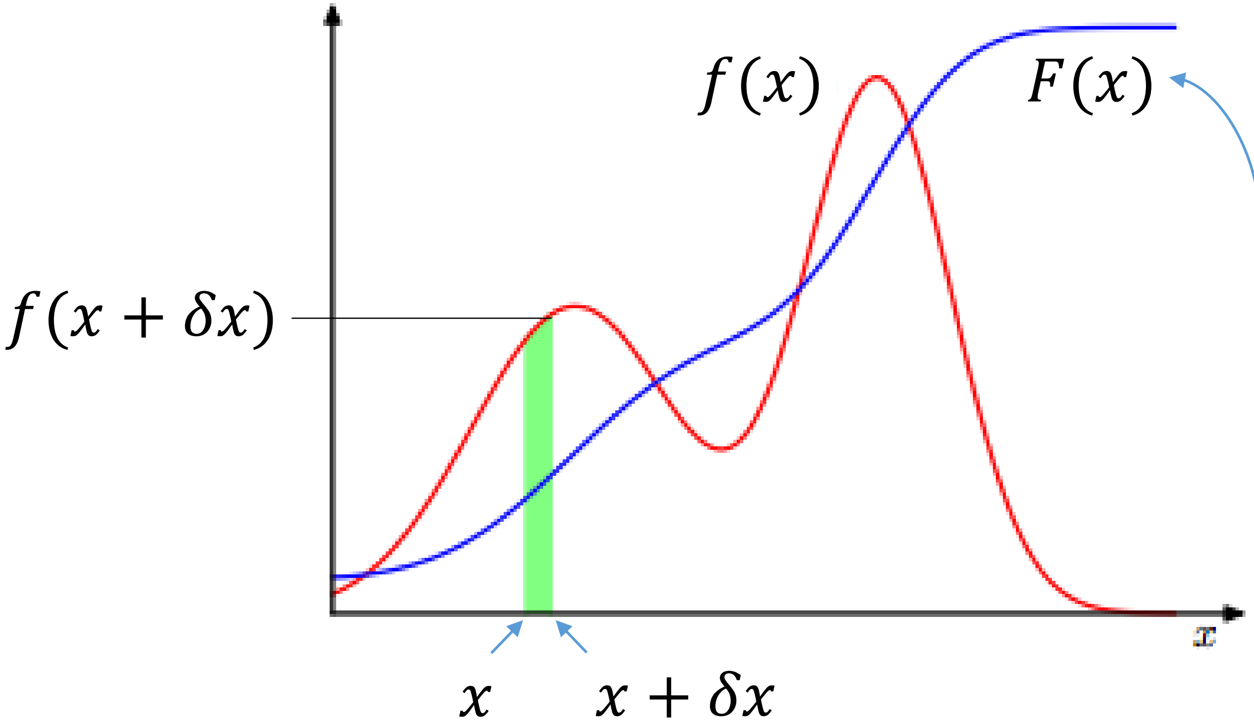
빗금영역의 면적

$$P(a \leq x \leq b) = \frac{(b-a) \cdot 1}{1} = \int_a^b f(x)dx$$

전체면적

$$P(-\infty \leq x \leq \infty) = \int_{-\infty}^{\infty} f(x)dx = 1$$

몇 가지 함수



$$P(a \leq x \leq b) = \int_a^b f(x)dx$$

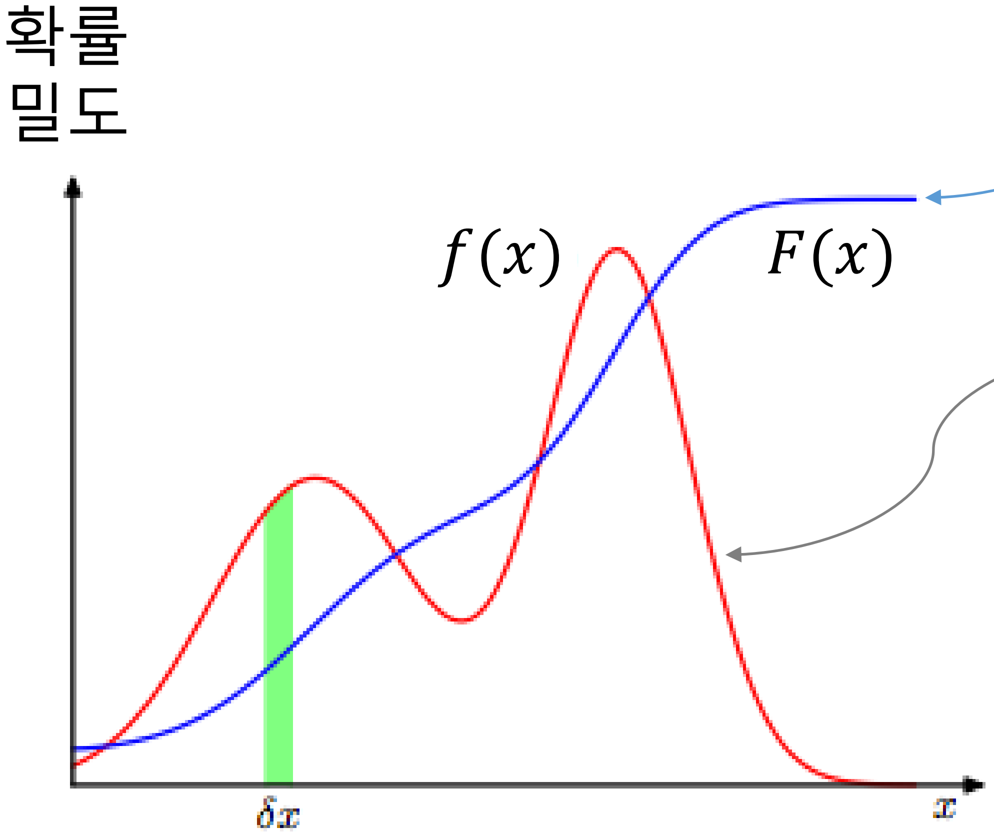
$$f(x) \geq 0$$

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

$$F(a) = \int_{-\infty}^{a} f(x)dx$$

$$P(X \leq a) = F(a)$$

$$\frac{d}{dx}F(x) = f(x)$$



누적확률밀도함수

누적확률분포함수

cumulative distribution function

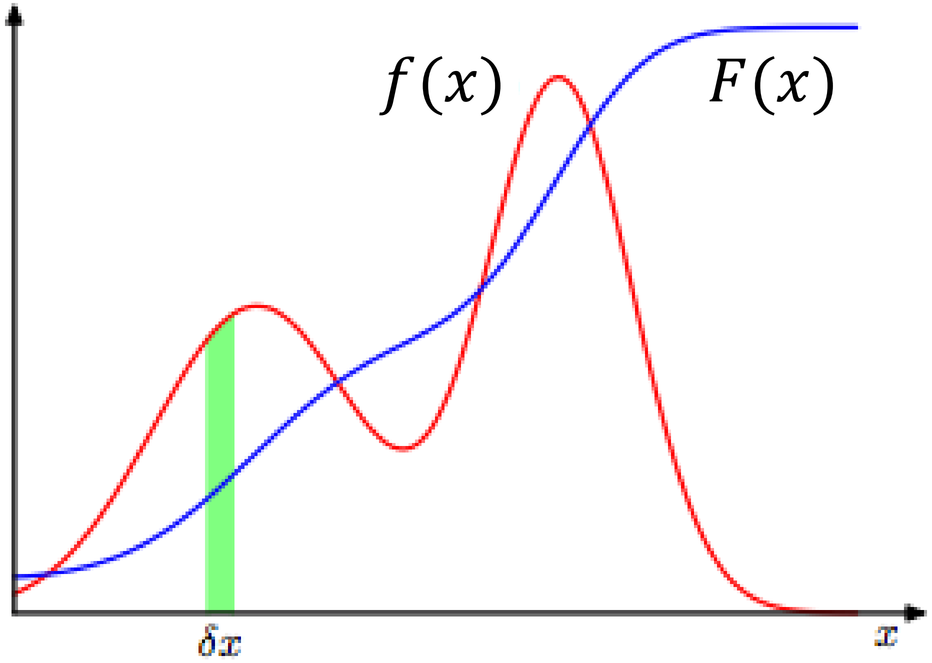
확률밀도함수

probability density function
(pdf)

$$\frac{\text{질량}}{\text{부피}} = \text{밀도}$$

부피×밀도=질량=확률

밀도



점, 선, 면, 부피, ...
scalar, linear, plane, volume, ...

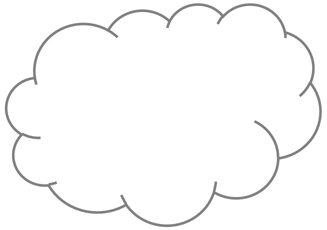
discrete 일 때,
확률질량함수
probability mass function
(pmf)

continuous 일 때,
확률밀도함수
probability density function
(pdf)

확률분포함수

- PDF (PMF)
- cumulative distribution function (CDF)

- 확률분포함수
- 확률분포



이거 확률분포가 어떻게 되?

pdf, pmf 가 뭐야?

확률변수에 대한 pdf가 뭐야?

$P(x)$

$P(X=x) \geq 0$

$\sum_x P(X=x) = 1$

$f(x)$

$f(x) \geq 0$

$\int f(x)dx = 1$

$P(a \leq X \leq b) = \int_a^b f(x)dx$

$P(X=0) = \frac{1}{\infty} = 0$

$f(X=x)$

$f_X(x)$

Joint Probability Density Function

- 확률변수 X, Y에 대한 분포함수; 결합분포함수; 2변수 이상

- $f(x, y) \geq 0$
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dxdy = 1$

Marginal Probability Density Function

- $f(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$
- $f(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$

Conditional PDF

$$p(y|x) = \frac{p(x,y)}{p(x)}$$

$$p(x,y) = p(y|x)p(x)$$

$$p(x,y,z) = p(x)p(y|x)p(z|x,y)$$

확률변수의 독립 조건

$$p(x, y) = p(x)p(y|x) = p(x)p(y)$$

기댓값

가중평균

x_1	x_2	x_3	x_4	x_5
1	2	3	4	5
n_1	n_2	n_3	n_4	n_5
1	6	1	1	1

$$\frac{1}{10} \times 1 + \frac{6}{10} \times 2 + \frac{1}{10} \times 3 + \frac{1}{10} \times 4 + \frac{1}{10} \times 5 = \frac{25}{10} = 2.5$$

$$\sum_{i=1}^{k} \frac{n_i}{N} x_i$$

k 는 서로 다른 숫자의 갯수

$N = \sum_{i=1}^{k} n_i$

$$\sum_{i=1}^{k} = \frac{n_i}{N} x_i$$

$$p(x_i) = \frac{n_i}{N}$$

$$\sum_{i=1}^{k} p(x_i) x_i$$

$$\boxed{\sum_{x} x \cdot p(x)}$$

이 확률을
정확하게 안다면,

$N = \infty$

동전을 10번 던지면
앞면이 나올 확률이 ½ ?

이상적인 평균이다.
기호와 이름을 하사 할 만 하다

이산확률변수의 기댓값

$$\mu = E[X] = \sum_{x} x \cdot p(x)$$

Expectation

확률질량함수

$$\frac{1}{n}\sum x = 1 \cdot \frac{1}{n} + 2 \cdot \frac{1}{n} + \cdots + n \cdot \frac{1}{n}$$

이산확률변수의 분산

- 평균에서 떨어져 있는 거리의 제곱의 기댓값

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

$$Var(X)$$

$$\sigma^2$$

연속확률변수의 기댓값, 분산

$$\mu = E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

$$\sigma^2 = Var(X) = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx$$

확률변수에 관한 기댓값 계산

- 확률변수 X : 동전 1개 던져서 앞면은 0 뒷면은 1

x	0	1
$P(X=x)$	$P(X=0)=0.5$	$P(X=1)=0.5$

$$E[X]=0\times p(0)+1\times p(1)=0.5$$

- 확률변수 W : 동전 앞면 4달러 받고, 뒷면 6달러 낸다.

w	4	-6
$P(W=w)$	$P(\mathrm{W}=4)=0.5$	$P(\mathrm{W}=-6)=0.5$

$$E[W]=-6\times p(0)+4\times p(1)=-1$$

$$W=10X-6 \qquad E[W]=E[10X-6]=10E[X]-6=-1$$

X에 관한 함수의 기댓값

- $E[g(X)] = \sum_x g(x)P(x)$

  pmf

- $E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$

  pdf

$P(X=x)$	$f(X=x)$
$P_X(x)$	$f_X(x)$

X, Y에 관한 함수의 기댓값

- $E[g(X, Y)] =$

기댓값의 성질

- $E[aX+b] = aE[X] + b$, $E[b] = b$
- $E[X+Y] = E[X] + E[Y]$

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

- $Var(X) = E[X^2] - \mu^2$
- $Var(aX+b) = a^2 Var(X)$
- $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$

- X와 Y가 독립이면 $\mathrm{E[XY] = E[X]E[Y]}, Cov(X,Y) = 0$

$E[aX+b]$

$$E[aX+b]=\sum_x (ax+b)P(x)$$

$$=a\sum_x xP(x)+b\sum_x P(x)$$

$$=aE[X]+b$$

상수는 앞으로 튀어 나온다.

$$E[b]=b$$

$$E[0X+b]=\sum_x (0x+b)P(x)$$

$$=0\sum_x xP(x)+b\sum_x P(x)$$

$$=0E[X]+b$$

$E[X+Y]$

$$E[X+Y] = \sum_x \sum_y (x+y) P(X=x, Y=y)$$

$$= \sum_x \sum_y x P(X=x, Y=y) + \sum_x \sum_y y P(X=x, Y=y)$$

marginalize

$$= \sum_x x P(X=x) + \sum_y y P(Y=y)$$

$$= E[X+Y] = E[X] + E[Y]$$

$E[X+Y]$

$$E[X+Y] = \iint (x+y) f_{X,Y}(x,y)\,dx\,dy$$

$$= \iint x f_{X,Y}(x,y)\,dx\,dy + \iint y f_{X,Y}(x,y)\,dx\,dy$$

$$= \int x f_X(x)\,dx + \int y f_Y(y)\,dy$$

$$= E[X] + E[Y]$$

marginalize

$$f_X(x) = \int f_{X,Y}(x,y)\,dy$$

$$f_Y(y) = \int f_{X,Y}(x,y)\,dx$$

$E[\ ]$: 하나의 연산자/함수

어떤 특징이 있지?

- $E[2] =$
- $E[\,2X + 3Y] =$
- $E[\,2X + 3Y + 4] =$

a=2+3;
a=add(2, 3)

함수 $\approx$ 연산자

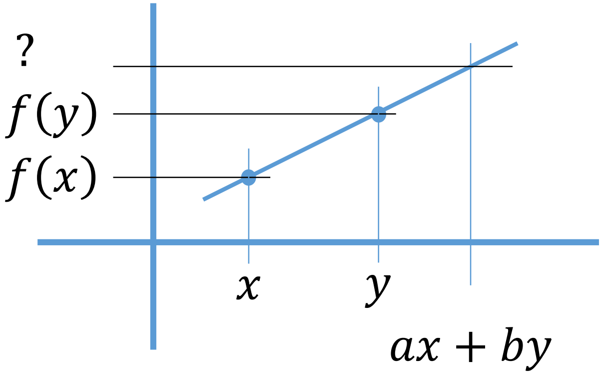
$f(x)$ 선형함수; linear function

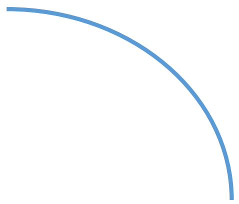
- $f[x]$, discrete
- $f(x)$, continuous
- $f(ax + by) = af(x) + bf(y)$

$$f(x_1 + x_2 + x_3) = f(x_1) + f(x_2) + f(x_3)$$

$$f\left(\sum x_i\right) = \sum f(x_i)$$

덧셈연산자와 **선형함수** 적용 순서를 바꿀 수 있다.





Taylor Series

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \frac{f^{(3)}(x)}{3!}h^3 + \cdots + \frac{f^{(n)}(x)}{n!}h^n + \cdots$$

$$\boxed{f(x+h) \approx f(x) + f'(x)h} + \frac{f''(x)}{2!}h^2 + \frac{f^{(3)}(x)}{3!}h^3 + \cdots + \frac{f^{(n)}(x)}{n!}h^n + \cdots$$

$Var(X)$

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

$$= E[X^2 - 2\mu X + \mu^2]$$

$$= E[X^2] - E[2\mu X] + \mathrm{E}[\mu^2]$$

$$= E[X^2] - 2\mu E[X] + \mathrm{E}[\mu^2]$$

$$= E[X^2] - 2\mu^2 + \mu^2$$

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

$$= \boxed{E[X^2] - E[X]^{2}}$$

$Var(aX + b)$

$$Var(aX + b) = E[(aX + b - E[aX + b])^2]$$

$$= E[(aX + b - au - b)^2]$$

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

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

$$= a^2Var(X)$$

$$\begin{aligned} \mathrm{Var}(b) \quad &= E[(b - E[b])^2] \\ &= E[(b - b)^2] \\ &= E[0] \\ &= 0 \end{aligned}$$

$Cov(X, Y)$; Covariance; 공분산;

$Cov(X, Y) =$

$$Cov(X+Z,Y) \qquad (X+Y)Z = XZ+YZ$$

$$Cov(X+Z,Y)$$

$$= E[(X+Z)Y] - E[X+Z]E[Y]$$

$$= E[XY] + E[ZY] - (E[X]+E[Z])E[Y]$$

$$= E[XY] - E[X]E[Y] + E[ZY] - E[Z]E[Y]$$

$$= Cov(X,Y) + Cov(Z,Y)$$

$$\underbrace{Cov\left(\sum}_{\text{이 두 연산자는 순서를 바꿀 수 있구나!}} X_i, Y\right) = \sum Cov(X_i, Y) \Rightarrow Cov\left(\sum_i X_i, \sum_j Y_j\right) =$$

$$Cov(X + Z, Y + W) \qquad (X+Z)(Y+W) = XZ + XY + ZY + ZW$$

$$
\begin{aligned}
&Cov(X + Z, Y) \\
&= E[(X+Z)Y] - E[X+Z]E[Y] \\
&= E[XY] + E[ZY] - (E[X] + E[Z])E[Y] \\
&= E[XY] - E[X]E[Y] + E[ZY] - E[Z]E[Y] \\
&= Cov(X, Y) + Cov(Z, Y)
\end{aligned}
$$

$$Cov\left(\sum X_i, Y\right) = \sum Cov(X_i, Y) \quad \Rightarrow \quad \boxed{Cov\left(\sum_i X_i, \sum_j Y_j\right) = \sum_i \sum_j Cov(X_i, Y_i)}$$

이 두 연산자는 순서를 바꿀 수 있구나!

$Var(X+Y)$

$$Var(X+Y)$$
$$= Cov(X+Y, X+Y)$$
$$= Cov(X,X) + Cov(X,Y) + Cov(Y,X) + Cov(Y,Y)$$
$$= Var(X) + Var(Y) + 2Cov(X,Y)$$

$$\boxed{Cov\left(\sum_i X_i, \sum_j Y_j\right) = \sum_i \sum_j Cov(X_i, Y_i)}$$

X, Y가 독립

$$E[XY]$$

$$= \sum_j \sum_i x_i y_j P(x_i, y_j)$$

$$= \sum_j \sum_i x_i y_j P(x_i) P(y_j)$$

$$= \sum_j y_j P(y_j) \sum_i x_i P(x_i)$$

$$= E[X]E[Y]$$

$$\boxed{E[XY] = E[X]E[Y]}$$

X, Y가 독립

$Cov(X, Y)$

$= E[XY] - E[X]E[Y]$

$= 0$

$$\boxed{Cov(X, Y) = 0}$$

참고문헌

- Sheldon M. Ross (이광수 역), Introduction to Probability and Statistics for Engineering and Scientists 5th Edition (이공계용 확률과 통계 5판), Academic Press (홍릉과학출판사)-CHAP.4