Foundations of Machine Learning
AI2000 and AI5000

FoML-03
Probability - Expectation, Variance and Gaussian Distribution

Dr. Konda Reddy Mopuri
Department of AI, IIT Hyderabad
July-Nov 2025

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

DiL
Data-driven Intelligence & Learning Lab

So far in FoML

- What is ML and the learning paradigms
- Probability refresher
    - Sum rule, product rule, Random variables, Bayes Theorem, Independence

Expectation, Variance and the Gaussian Distribution

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Data-driven Intelligence
& Learning Lab

Expectation

- Random variable X and a function $f: X \to \mathbb{R}$

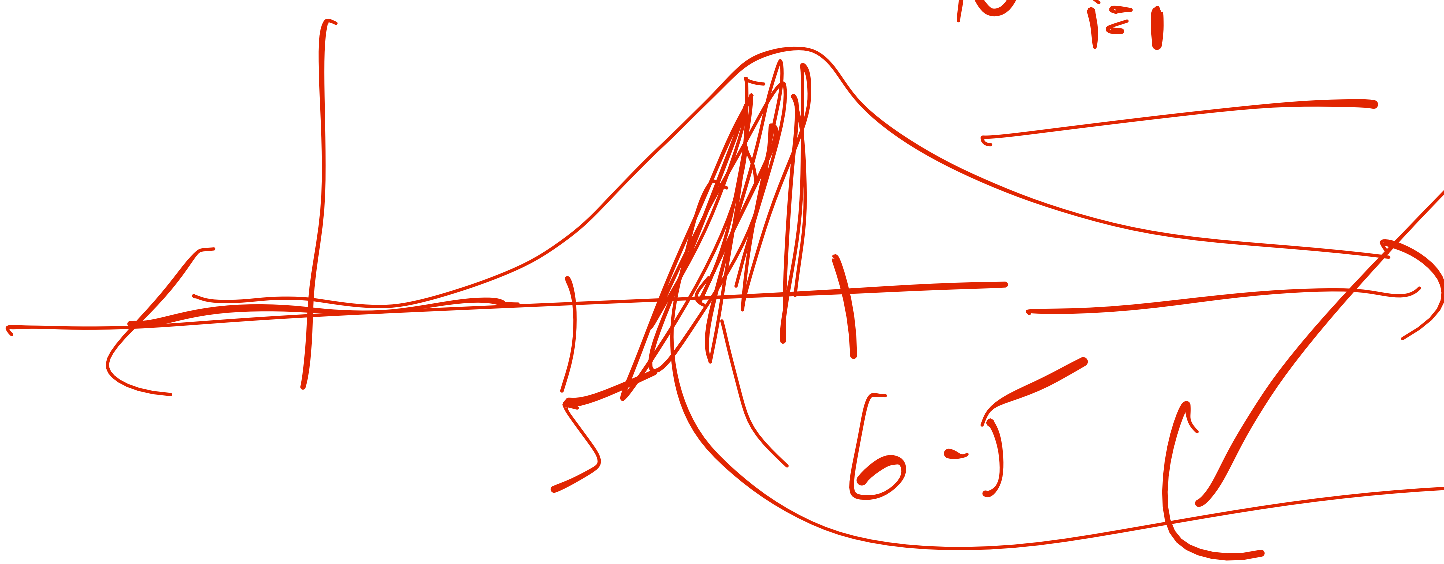
$$\mathbb{E}[f] = \mathbb{E}_{x \sim p(X)}[f(x)] = \int_X f(x)\, p(x)\, dx$$

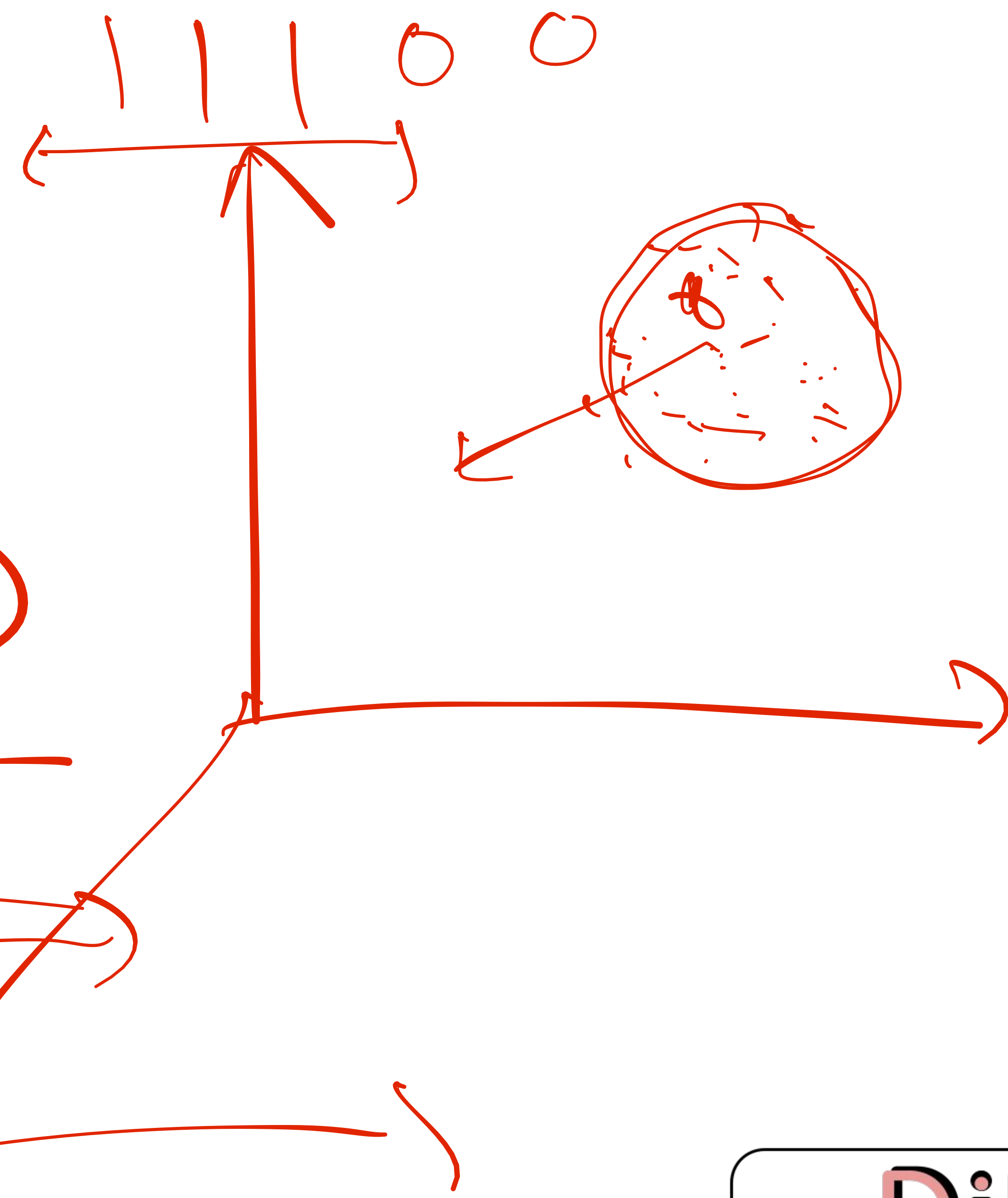
$$\sum_X f(x)\, p(x)$$

Expectation

- For N points drawn from p(X)

$$\mathbb{E}[f] = \frac{1}{N}\sum_{i=1}^{N} f(x_i)$$





Expectation

- Conditional expectation

$$\mathbb{E}[f/y] = \mathbb{E}_{x \sim p(X/Y=y)}[f(x)] = \sum_{x \in A} f(x)\, P(X=x \mid Y=y)$$

$$f(\alpha X + \beta Y) = \alpha f(X) + \beta f(Y)$$

Variance

- Expected quadratic distance between f and its mean $\mathbb{E}[f]$

$$\mathrm{var}(f) = \mathbb{E}_{x\sim P(x)}\left(f(x) - \mathbb{E}[f(x)]\right)^2 \qquad \left(x - \mathbb{E}[x]\right)^2$$

$$= \mathbb{E}[f(x)^2] - \mathbb{E}[f(x)]^2$$

Covariance

- Measures the extent to which two random variables X and Y vary together

$$\mathrm{cov}[X,Y] = E\left[(X - E(X))(Y - E(Y))\right]$$

$$= \underbrace{E[XY]}_{P(x,y)} - \underbrace{E[X]}_{P(x)}\, \underbrace{E[Y]}_{P(y)}$$

$P(x,y)$

Covariance

$D \times 1$

- X and Y are vectors of random variables
- Covariance matrix

$$\mathrm{cov}[X,Y] = E_{P(x,y)}\left[X - E(X)\right]\left[Y - E(Y)\right]^T$$

(i, j)

$D \times D$

Covariance

- Between independent variables

$$\mathrm{cov}[X,Y] = E[xy] - E[x]E[y]$$

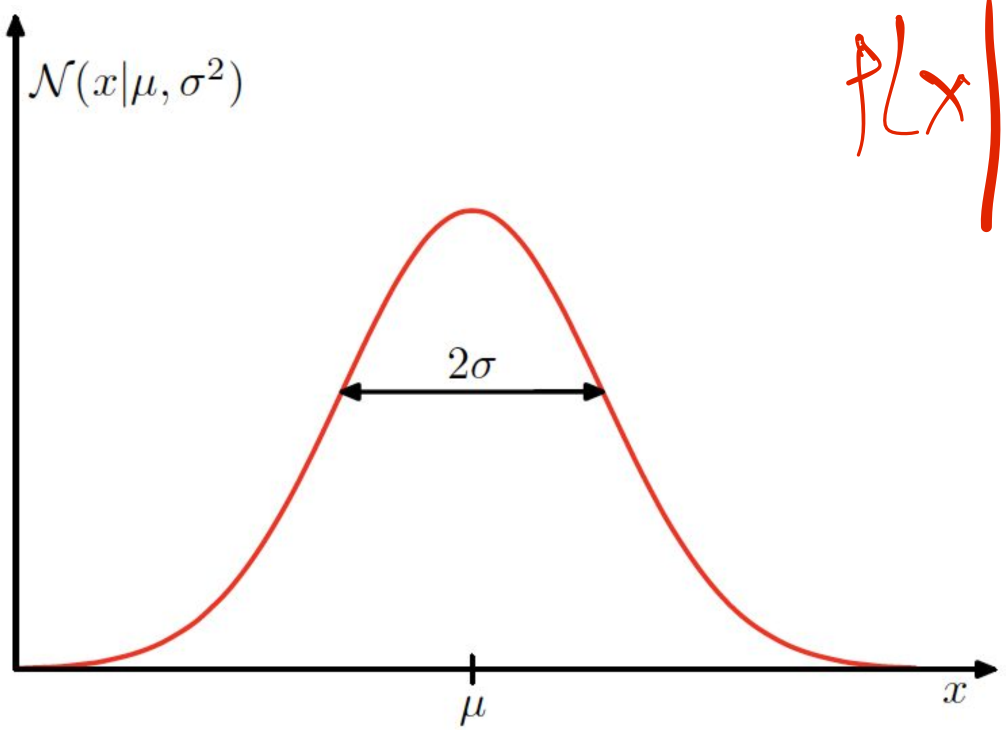
$$= E(x)E(y) -$$

$$= 0$$

Gaussian Distribution

Gaussian Distribution

$x \sim \mathcal{N}(\mu, \sigma^2)$ μ σ



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

$$\mathcal{N}(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

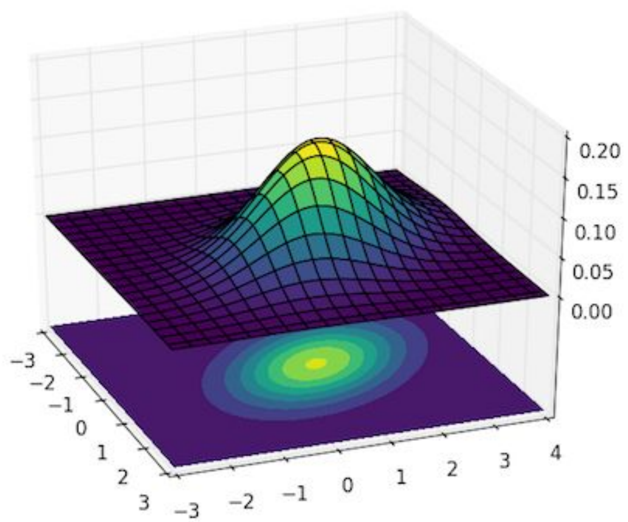
$$x \sim \mathcal{N}(x|\mu,\sigma^2)$$

$$\mathbb{E}[x] = \mu \qquad \mathrm{Var}(x) = \sigma^2$$

Multivariate Gaussian Distribution

- D-dimensional vector $\mathbf{x} = (x_1, x_2, \ldots x_D)^T$

$$\mathcal{N}(\mathbf{x}|\mu, \Sigma) = \frac{1}{(2\pi)^{D/2}|\mathbf{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$



Next

Maximum Likelihood Principle