Machine Learning Foundations

Chapter 6: Probability

Harish Guruprasad Ramaswamy
IIT Madras

Outline for Chapter 6 : Probability

6.1 : Discrete Random Variables

6.2 : Continuous Random Variables

6.3 : Maximum Likelihood and other advanced topics

Outline for Chapter 6 : Probability

6.1 : Discrete Random Variables

1. Probability space
2. Conditioning
3. Random variables
4. Expectation and Variance
5. Multiple Random Variables
6. **Bernoulli, Binomial, Poisson and Geometric RVs**

6.2 : Continuous Random Variables

6.3 : Maximum Likelihood and other advanced topics

Bernoulli distribution

$$f(1) = p$$

$$f(0) = 1-p$$

$$X \sim \text{Bernoulli}(p)$$

$$E[X] = P(X=1)$$

$$= p$$

$$\text{Var}[X] = EX^2 - (EX)^2$$

$$= p - p^2$$

$$= p(1-p)$$

Bernoulli distribution

Binomial distribution

$Bin(n, p)$

$X_1, X_2, \ldots, X_n$ are independent Bernoulli (p) RVs.

$$X = \sum_{i=1}^{n} X_i$$

$$X \sim Bin(n, p)$$

HHTTT

TTHHT

$$f_X(k) = P(X = k)$$

$$= \left({}^{n}C_{k} \right) \cdot (p)^{k} (1-p)^{n-k}$$

Binomial distribution

$$\sum_{k=0}^{n} f(k) = \sum_{k=0}^{n} \left({}^{n}C_{k}\right) \cdot (p)^{k} (1-p)^{n-k}$$

$$= \left(p + (1-p)\right)^{n}$$

$$= 1$$

$$EX = ?$$

$$= \sum_{i=1}^{n} EX_i = np$$

$$Var[X] = \sum_{i=1}^{n} var[X_i] = np(1-p)$$

$$= npq$$

Poisson distribution

$\text{Bin}(n, p)$

$n \downarrow$ Large, $p \downarrow$ Small

$X \sim \text{Poisson}(\lambda)$

$$f_X(k) = P(X = k)$$

$$= e^{-\lambda} \frac{\lambda^k}{k!}$$

$$\sum_{k=0}^{\infty} f_X(k) = e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$

$$= e^{-\lambda} \cdot e^{\lambda} = 1$$

Poisson distribution

$X \sim \text{Poisson}(\lambda)$ takes value $0, 1, \ldots$

$$EX = \sum_{k=0}^{\infty} k \cdot f_X(k)$$

$$= \sum_{k=1}^{\infty} k \cdot e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

$$= e^{-\lambda} \cdot \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!}$$

$$= e^{-\lambda} \cdot \lambda \cdot \sum_{k=1}^{\infty} \frac{\lambda^{(k-1)}}{(k-1)!} = \lambda$$

Geometric distribution

$X \sim \text{Geom}(p)$ X takes values $1, \ldots$

$$f_X(k) = P(X = k)$$

$$= (1-p)^{k-1} p$$

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

$$= p \cdot \frac{1}{1-(1-p)} = 1$$

Geometric distribution

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

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

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

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

$$EX = \frac{1}{p}$$