Lecture Notes

on

Information Theory and Coding

GALGOTIAS UNIVERSITY
(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

July 2020
(Be safe and stay at home)

Independence between events

- N events are statistically independent if the intersection of the events contained in any subset of those N events have probability equal to the product of the individual probabilities
- Example: Three events A, B and C are independent if:

$$P(A \cap B) = P(A)P(B), \; P(A \cap C) = P(A)P(C), \; P(B \cap C) = P(B)P(C)$$

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

Random Variables

- A random variable (rv) is a function that maps each $\omega \in \Omega$ to a real number.

$$\begin{aligned} X \quad : \quad & \Omega \to \mathbb{R} \\ & \omega \to X(\omega) \end{aligned}$$

- Through a random variable, subsets of Ω are mapped as subsets (intervals) of the real numbers.

$$P(X \in I) = P(\{w | X(\omega) \in I\})$$

Random Variables

- A real random variable is a function whose domain is Ω and such that
  - for all real number x, the set $A_x = \{\omega | X(\omega) \leq x\}$ is an event.
  - $P(w|X(w) = \pm\infty) = 0$.

Cumulative Distribution Function

$$F_X : \mathbb{R} \to [0,1]$$
$$X \to F_X(x) = P(X \le x) = P(\omega | X(\omega) \le x)$$

- $F_X(\infty) = 1$
- $F_X(-\infty) = 0$
- If $x_1 < x_2$, $F_X(x_2) \ge F_X(x_1)$.
- $F_X(x^+) = \lim_{\epsilon \to 0} F_X(x+\epsilon) = F_X(x)$. (continuous on the right side).
- $F_X(x) - F_X(x^-) = P(X = x)$.

Types of Random Variables

- Discrete: Cumulative function is a step function (sum of unit step functions)

$$F_X(x) = \sum_i P(X = x_i) u(x - x_i)$$

  where $u(x)$ is the unit step funtion.
- Example: X is the random variable that describes the outcome of the roll of a die. $X \in \{1, 2, 3, 4, 5, 6\}$

Types of Random Variable

- Continous: Cumulative function is a continous function.
- Mixed: Neither discrete nor continous.

Probability Density Function

- It is the derivative of the cumulative distribution function:

$$p_X(x) = \frac{d}{dx} F_X(x)$$

- $\int_{-\infty}^{x} p_X(x)dx = F_X(x)$.
- $p_X(x) \geq 0$.
- $\int_{-\infty}^{\infty} p_X(x)dx = 1$.
- $\int_{a}^{b} p_X(x)dx = F_X(b) - F_X(a) = P(a \leq X \leq b)$.
- $P(X \in I) = \int_{I} p_X(x)dx,\ I \subset \mathbb{R}$.

Discrete Random Variables

- Let us now focus only on discrete random variables.
- Let X be a random variable with sample space $\mathcal{X}$
- The probability mass function (probability distribution function) of X is a mapping $p_X(x) : \mathcal{X} \to [0, 1]$ satisfying:

$$\sum_{X \in \mathcal{X}} p_X(x) = 1$$

- The number $p_X(x) := P(X = x)$

Discrete Random Vectors

- Let $Z = [X, Y]$ be a random vector with sample space $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$
- The joint probability mass function (probability distribution function) of Z is a mapping $p_Z(z) : \mathcal{Z} \to [0, 1]$ satisfying:

$$\sum_{Z \in \mathcal{Z}} p_Z(z) = \sum_{x,y \times \mathcal{Y}} p_{XY}(x, y) = 1$$

- The number $p_Z(z) := p_{XY}(x, y) = P(Z = z) = P(X = x, Y = y)$.

Discrete Random Vectors

- Marginal Distributions

$$p_X(x) = \sum_{y \in \mathcal{Y}} p_{XY}(x, y)$$

$$p_Y(y) = \sum_{x \in \mathcal{X}} p_{XY}(x, y)$$

Discrete Random Vectors

- Conditional Distributions

$$p_{X|Y=y}(x) = \frac{p_{XY}(x,y)}{p_Y(y)}$$

$$p_{Y|X=x}(y) = \frac{p_{XY}(x,y)}{p_X(x)}$$

Discrete Random Vectors

- Random variables X and Y are independent if and only if

$$p_{XY}(x, y) = p_X(x)p_Y(y)$$

- Consequences:

$$p_{X|Y=y}(x) = p_X(x)$$

$$p_{Y|X=x}(y) = p_Y(y)$$

Moments of a Discrete Random Variable

- The $n-$th order moment of a discrete random variable X is defined as:

$$E[X^n] = \sum_{x \in \mathcal{X}} x^n p_X(x)$$

- if $n = 1$, we have the mean of X, $m_X = E[X]$.
- The $m-$th order central moment of a discrete random variable X is defined as:

$$E[(X - m_X)^m] = \sum_{x \in \mathcal{X}} (x - m_X)^m p_X(x)$$

- if $m = 2$, we have the variance of X, σ_X^2.

Moments of a Discrete Random Vector

- The joint moment $n-$th order with relation to X and $k-$th order with relation to Y:

$$m_{nk} = E[X^n Y^k] = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} x^n y^k p_{XY}(x, y)$$

- The joint central $n-$th order with relation to X and $k-$th order with relation to Y:

$$\mu_{nk} = E[(X - m_X)^n (Y - m_Y)^k] == \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} (x - m_X)^n (y - m_Y)^k p_{XY}(x, y)$$

Correlation and Covariance

- The correlation of two random variables X and Y is the expected value of their product (joint moment of order 1 in X and order 1 in Y):

$$Corr(X, Y) = m_{11} = E[XY]$$

- The covariance of two random variables X and Y is the joint central moment of order 1 in X and order 1 in Y:

$$Cov(X, Y) = \mu_{11} = E[(X - m_X)(Y - m_Y)]$$

- $Cov(X, Y) = Corr(X, Y) - m_X m_Y$
- Correlation Coefficient:

$$\rho_{XY} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} \quad \rightarrow \quad -1 \leq \rho_{XY} \leq 1$$