Lecture Summary: Joint PMFs for More Than Two Random Variables

Source: Lec1.5.pdf

Key Points

- **Introduction to Multi-Variable Joint PMFs:**
  - Extending joint PMFs to more than two random variables is conceptually straightforward.
  - For discrete random variables $X_1, X_2, \ldots, X_n$, the joint PMF is defined as:

$$f_{X_1,X_2,\ldots,X_n}(t_1, t_2, \ldots, t_n) = P(X_1 = t_1, X_2 = t_2, \ldots, X_n = t_n).$$

  - It is a function mapping the Cartesian product of the ranges of $X_1, X_2, \ldots, X_n$ to probabilities.
- **Example 1: Coin Tosses:**
  - Experiment: Tossing a fair coin three times, defining random variables X_1, X_2, X_3 to represent outcomes of each toss.
  - Possible outcomes: $\{000, 001, 010, 011, 100, 101, 110, 111\}$ (binary representation).
  - Joint PMF:

$$f_{X_1,X_2,X_3}(t_1, t_2, t_3) = \frac{1}{8}, \quad \text{for all } (t_1, t_2, t_3).$$

  - Uniform distribution across all outcomes.
- **Example 2: Three-Digit Random Numbers:**
  - Experiment: Randomly generate a three-digit number (000 to 999).
  - Define random variables:
    - X_1: First digit (hundreds place).
    - X_2: Last digit (units place).
    - X_3: Modulo 2 of the number (even or odd).
  - Example joint PMF:
    - $f_{X_1,X_2,X_3}(0, 0, 0) = \frac{10}{1000} = \frac{1}{100}$ (10 possibilities: numbers like 000, 020, etc.).
    - $f_{X_1,X_2,X_3}(1, 1, 1) = \frac{1}{100}$ (numbers like 111, 131, etc.).
  - Some combinations may have zero probability (e.g., $X_3 = 0$ and $X_2 = 1$ is impossible).
- **Example 3: Powerplay Overs in IPL Cricket:**
  - Experiment: Analyze runs scored in six deliveries of a powerplay over.
  - Random variables: $X_1, X_2, \ldots, X_6$, representing runs scored on each ball.
  - Range: $\{0, 1, 2, 3, 4, 5, 6, 7, 8\}$ for each X_i (7 for no-ball with six hit, 8 for rare cases like over-throws).

- Joint PMF:
$$f_{X_1,X_2,\ldots,X_6}(t_1,t_2,\ldots,t_6),$$
represents probabilities for sequences like $(0,1,4,2,0,0)$.
- Complexity: With $9^6 = 531441$ possibilities, writing down the entire PMF is impractical.

- **Challenges and Practical Considerations:**
  - As the number of variables increases, enumerating all joint probabilities becomes infeasible.
  - In modern data problems, joint PMFs are often too vast to define explicitly.
  - Marginalization and conditional PMFs are practical tools for analyzing complex scenarios.

Simplified Explanation

What Are Multi-Variable Joint PMFs? They extend joint PMFs to more than two random variables, defining probabilities for all combinations of outcomes.

Examples: 1. Coin tosses: Uniform probabilities across all combinations. 2. IPL powerplay: Joint probabilities for runs scored on each ball are complex and computationally intensive.

Why Use Marginals and Conditionals? For high-dimensional scenarios, joint PMFs are unwieldy. Marginalization and conditioning simplify analysis by focusing on subsets of variables.

Conclusion

In this lecture, we:

- Defined joint PMFs for more than two random variables.
- Explored examples ranging from simple (coin tosses) to complex (IPL cricket).
- Highlighted challenges and the importance of using marginal and conditional distributions for practical analysis.

These tools are critical for tackling real-world problems involving multiple random variables.