Independence and Random Variables

Lecture 3

Grading

- Weekly Assignments: 20%
  - 8 Tests, 2 Per Quarter 5 marks and 30 mins
  - 2-3 problems variants randomly chosen from selected problems
    - 6 Problems per week: 3 unsolved, 3 from OCW Recitations
  - Marks of best 4 used for grading
- Quiz 1, Quiz 2: 10% each
  - Weight: 10% each
- MidSem: 25%
- EndSem: 35%
- Bonus Marks: 5%
  - If you write 7/8 tests, the average score among the tests will be added to the total score.
  - -

Reading

Read: [BT] Section 1.5, 2.1-2.4. [OCW] Lecture 3, 5.

Explore: [WF] Chapter 5, Section 3, 4. Chapter 9 Sections 1-4.

[BT] Introduction to Probability, 2nd Edition by Dimitri P. Bertsekas and John N. Tsitsiklis.

[WF] An Introduction to Probability Theory and Its Applications, Volume 1 by William Feller.

[OCW] Probabilistic Systems Analysis and Applied Probability Online Resource. MIT OCW https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-041sc-probabilistic-systems-analysis-and-applied-probability-fall-2013/index.htm

Previously

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}, \qquad \text{assuming } \mathbf{P}(B) > 0$$

- Multiplication rule:

$$\mathbf{P}(A \cap B) = \mathbf{P}(B) \cdot \mathbf{P}(A \mid B) = \mathbf{P}(A) \cdot \mathbf{P}(B \mid A)$$

- Total probability theorem:

$$\mathbf{P}(B) = \mathbf{P}(A)\mathbf{P}(B \mid A) + \mathbf{P}(A^c)\mathbf{P}(B \mid A^c)$$

- Bayes rule:

$$\mathbf{P}(A_i \mid B) = \frac{\mathbf{P}(A_i)\mathbf{P}(B \mid A_i)}{\mathbf{P}(B)}$$

Today & Upcoming

1. Independent Events
   a. For a Pair
   b. For a Collection

2. Random Variable
   a. Prob. Mass. Functions
   b. Mean
   c. Expectation

- Thursday Class & Friday Tutorial:
  - Problem Solving
  - Doubt Clearing
  - Bonus Topic

- Next Week
  - No Classes (Mon and Thu Holidays)
  - Test 2 on Friday (16th August)
    - Based on 6 + 9 problems
    - 9 new problems will be shared by end of this week.

Independence of two events

- **“Defn:”** $\mathbf{P}(B \mid A) = \mathbf{P}(B)$
  - “occurrence of A provides no information about B's occurrence”
- Recall that $\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B \mid A)$
- **Defn:** $\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$
- Symmetric with respect to A and B
  - applies even if $\mathbf{P}(A) = 0$
  - implies $\mathbf{P}(A \mid B) = \mathbf{P}(A)$

Examples:

1. A card is chosen at random from a deck of cards. The events “spade” and “ace”.

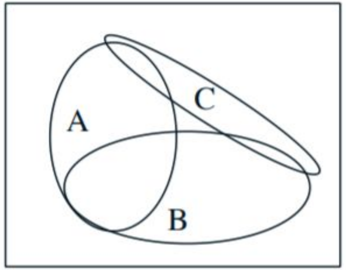
2. Random permutation of a,b,c,d
   P: a precedes b
   Q: c precedes d

3. Consider 3 children families.
   H: children of both sexes
   A: at most 1 girl
   What about these events in families of 4 children?

Conditioning Affects Independence

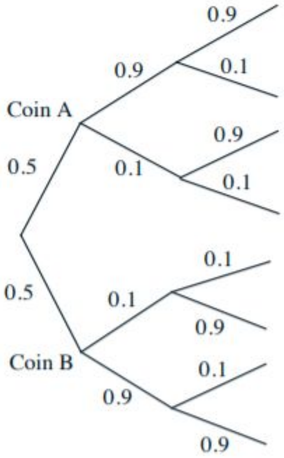
- Conditional independence, given C, is defined as independence under probability law $\mathbf{P}(\cdot \mid C)$

- Assume A and B are independent



- If we are told that C occurred, are A and B independent?

- Two unfair coins, A and B:
  $\mathbf{P}(H \mid \text{coin } A) = 0.9$, $\mathbf{P}(H \mid \text{coin } B) = 0.1$
  choose either coin with equal probability



- Once we know it is coin A, are tosses independent?
- If we do not know which coin it is, are tosses independent?
  - Compare:
    $\mathbf{P}(\text{toss } 11 = H)$
    $\mathbf{P}(\text{toss } 11 = H \mid \text{first 10 tosses are heads})$

Independence of a collection of events

- Intuitive definition:
  Information on some of the events tells us nothing about probabilities related to the remaining events
  - E.g.:

$$\mathbf{P}(A_1 \cap (A_2^c \cup A_3) \mid A_5 \cap A_6^c) = \mathbf{P}(A_1 \cap (A_2^c \cup A_3))$$

- Mathematical definition:
  Events $A_1, A_2, \ldots, A_n$
  are called **independent** if:

$$\mathbf{P}(A_i \cap A_j \cap \cdots \cap A_q) = \mathbf{P}(A_i)\mathbf{P}(A_j)\cdots\mathbf{P}(A_q)$$

  for any distinct indices $i, j, \ldots, q$, (chosen from $\{1, \ldots, n\}$)

Independence vs. pairwise independence

- Two independent fair coin tosses
  - A: First toss is H
  - B: Second toss is H
  - $\mathbf{P}(A) = \mathbf{P}(B) = 1/2$

HH	HT
TH	TT

  - C: First and second toss give same result
  - $\mathbf{P}(C) =$
  - $\mathbf{P}(C \cap A) =$
  - $\mathbf{P}(A \cap B \cap C) =$
  - $\mathbf{P}(C \mid A \cap B) =$
- Pairwise independence **does not** imply independence

The king's sibling

- The king comes from a family of two children. What is the probability that his sibling is female?

Random variables

- An assignment of a value (number) to every possible outcome
- Mathematically: A function from the sample space Ω to the real numbers
  - discrete or continuous values
- Can have several random variables defined on the same sample space
- Notation:
  - random variable X
  - numerical value x

Probability mass function (PMF)

- ("probability law", "probability distribution" of X)
- Notation:

$$\begin{aligned} p_X(x) &= \mathbf{P}(X = x) \\ &= \mathbf{P}(\{\omega \in \Omega \text{ s.t. } X(\omega) = x\}) \end{aligned}$$

- $p_X(x) \geq 0 \qquad \sum_x p_X(x) = 1$
- **Example:** X=number of coin tosses until first head
  - assume independent tosses, $\mathbf{P}(H) = p > 0$

$$\begin{aligned} p_X(k) &= \mathbf{P}(X = k) \\ &= \mathbf{P}(TT \cdots TH) \\ &= (1-p)^{k-1}p, \qquad k = 1, 2, \ldots \end{aligned}$$

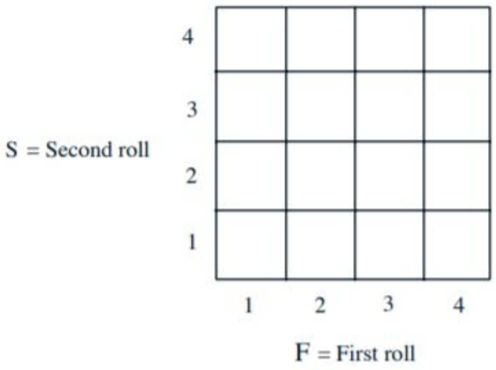
  - **geometric** PMF

How to compute a PMF $p_X(x)$

- collect all possible outcomes for which X is equal to x
- add their probabilities
- repeat for all x

- **Example:** Two independent rools of a fair tetrahedral die

  F: outcome of first throw
  S: outcome of second throw
  $X = \min(F, S)$



$p_X(2) =$

Binomial PMF

- X: number of heads in n independent coin tosses
- $\mathbf{P}(H) = p$
- Let $n = 4$

$$\begin{aligned} p_X(2) &= \mathbf{P}(HHTT) + \mathbf{P}(HTHT) + \mathbf{P}(HTTH) \\ &\quad + \mathbf{P}(THHT) + \mathbf{P}(THTH) + \mathbf{P}(TTHH) \\ &= 6p^2(1-p)^2 \\ &= \binom{4}{2} p^2 (1-p)^2 \end{aligned}$$

In general:

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \ldots, n$$

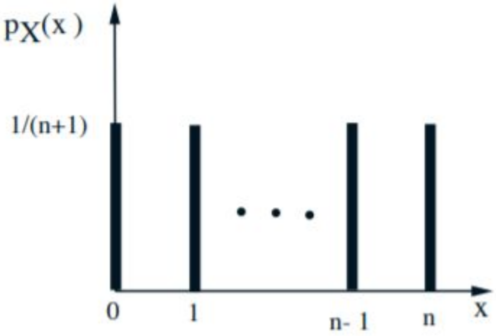
Expectation

- Definition:

$$\mathbf{E}[X] = \sum_x x p_X(x)$$

- Interpretations:
  - Center of gravity of PMF
  - Average in large number of repetitions of the experiment (to be substantiated later in this course)

- Example: Uniform on $0, 1, \dots, n$



$$\mathbf{E}[X] = 0\times\frac{1}{n+1} + 1\times\frac{1}{n+1} + \cdots + n\times\frac{1}{n+1} =$$

Properties of expectations

- Let X be a r.v. and let $Y = g(X)$
  - Hard: $\mathbf{E}[Y] = \sum_y y p_Y(y)$
  - Easy: $\mathbf{E}[Y] = \sum_x g(x) p_X(x)$

- Caution: In general, $\mathbf{E}[g(X)] \neq g(\mathbf{E}[X])$

Properties: If α, β are constants, then:

- $\mathbf{E}[\alpha] =$

- $\mathbf{E}[\alpha X] =$

- $\mathbf{E}[\alpha X + \beta] =$

Variance

Recall: $\mathbf{E}[g(X)] = \sum_x g(x) p_X(x)$

- **Second moment:** $\mathbf{E}[X^2] = \sum_x x^2 p_X(x)$
- **Variance**

$$\begin{aligned}\mathrm{var}(X) &= \mathbf{E}\left[(X - \mathbf{E}[X])^2\right] \\ &= \sum_x (x - \mathbf{E}[X])^2 p_X(x) \\ &= \mathbf{E}[X^2] - (\mathbf{E}[X])^2\end{aligned}$$

Properties:

- $\mathrm{var}(X) \geq 0$
- $\mathrm{var}(\alpha X + \beta) = \alpha^2 \mathrm{var}(X)$