Elements of Probability and Statistics

Lecture 02: Axiomatic Definition of Probability

IAI, TCG-CREST — August 11, 2023

2.1 (contd.) Applying elemental counting to calculate naive probabilities

	Order matters	Order does not matter
SWR	n^k	$\binom{n+k-1}{k}$
SWOR	$\frac{n!}{(n-k)!}$	$\binom{n}{k}$

Table 1: Counting over four cases: Sampling with replacement (SWR), sampling without replacement (SWOR), and for both situations considering whether the order of the objects in a sample matters or not.

Q: From a draw of 5 cards from a pack of 52 cards that are shuffled, what is the probability of getting a *full house* in poker? A full house contains three cards with the same face, while the other two cards are also of the same face. For example, a draw of 'KKK33' containing three Kings and two 3's is a full house.

Q: *Birthday Problem:* In a room with k individuals, what is the probability that at least two individuals have the same birthday? We assume the knowledge of one individual's birthday would not inform us about the birthday of another individual.

Q: *Newton-Pepys Problem:* Samuel Pepys once asked Isaac Newton the following problem related to gambling: Which of the following events had the highest probability: (i) at least one 6 appears when 6 dice are rolled; (ii) at least two 6 appear when 12 dice are rolled; or (iii) at least three 6 appear when 18 dice are rolled.

2.2 Axiomatic definition of Probability

Defn. A probability space consists of (i) a sample space S, and (ii) a probability function $P : S \to [0, 1]$ which takes as input any event $A \subseteq S$ and provides as output the probability of that event $P(A) \in [0, 1]$. The function P must satisfy the following axioms.

1. $P(\phi) = 0,\ P(S) = 1$.
2. If $A_1, A_2, ...$ are disjoint events, then,

$$P(\cup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} P(A_j).$$

These axioms define the probability function P, but how should P be interpreted is a question that leads to two schools of thoughts. The *frequentist* interpretation of probability is P is the long-run frequency over a large number of repetitions of the random experiment. The *Bayesian* interpretation is that P represents an assumption of a degree of belief about an event. In different scenarios, either one of the interpretations can be more helpful to understand the problem being studied.

Theorem 1 *For any two events A and B, the following are true.*

1. $P(A^c) = 1 - P(A)$.

2. *If* $A \subseteq B$*, then* $P(A) \leq P(B)$.

3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Theorem 2 *For any events* $A_1, ..., A_n$,

$$P(\cap_{i=1}^{\infty} A_i) = \sum_i P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - ... + (-1)^{n+1} P(A_1 \cap ... \cap A_n).$$

Q: *de Montmort's matching problem:* There are n cards, labeled 1 to n, that are shuffled and placed face down. Starting from the card in position 1, each card is sequentially flipped to reveal their number. If at any point, the card at position i is the card with number i on its face, you win the game. What is the probability of winning the game?

Avisek Gupta
Postdoctoral Fellow
IAI, TCG CREST.