STA257 Notes

Max Xu

'25 Fall

Contents

1 Day 1: Sample Spaces and Probability (Sept 3, 2025) 2

2 Day 2: (Sept 8, 2025) 3

§1 Day 1: Sample Spaces and Probability (Sept 3, 2025)

Life is very random and uncertain, with interesting problems to solve (e.g. whats the probability that I win the lottery). This course uses certain mathematics to study the uncertainty of probabilities. You will be able to solve problems like this:

Problem 1.1. Which is more likely: getting at least one six when rolling a fair 6 sided die 4 times, or getting one pair of sixes when rolling two six sided dice 24 times?

5% of your grade is poll-based, more info here.

Definition 1.2 (Sample Space)**.** A non-empty set containing all possible outcomes, written S.

e.g. coin-flipping: $S = \{\text{Heads}, \text{Tails}\}$, two die: $S = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$

Definition 1.3 (Event)**.** Any subset $A \subseteq S$ is an event.

Prof says in some continuous sample spaces, there may exist some non-measurable subsets to which the probability measure defined later won't work on, but don't worry about it in this course. (yay!!)

Definition 1.4 (Probability)**.** For any event A, define probability $\mathrm{P}(A)$ that satisfies:

- For all $A \subseteq S$, $0 \leq \mathrm{P}(A) \leq 1$
- $A = S$, $\mathrm{P}(A) = \mathrm{P}(S) = 1$
- $A = \emptyset$, corresponding to no outcome, then $\mathrm{P}(A) = \mathrm{P}(\emptyset) = 0$
- **Additivity**: if $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$

If $A_1, \ldots, A_n$ are disjoint[1] events, we have

- **Finite Additivity**: For some $n \in \mathbb{N}$,

$$\mathrm{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathrm{P}(A_i)$$

- **Countable Additivity**:

$$\mathrm{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathrm{P}(A_i)$$

[2]

When looking at the probability of getting heads from a coinflip, $\mathrm{P}(H)$ with $S = \{T, H\}$ is really shorthand for $\mathrm{P}(\{H\})$, since H may not be a subset of S. For uniformly picking any number between 0 and 1, denoted $\text{Uniform}[0, 1]$, we can define a probability $\mathrm{P}([a, b]) = b - a$ whenever $0 \leq a \leq b \leq 1$. (don't know what uniform means yet)

Note that by definition of probability, $\mathrm{P}(A_i)$ is positive, so the right hand side is an absolutely convergent series (prof didn't mention this).

[1] prof said this but i think he meant pairwise disjoint?

[2] we could've actually done this with just 3, see probability axioms

§2 Day 2: (Sept 8, 2025)

§2.1 Additional Properties of Probability

Today we will be deriving properties of probability from the 'axioms' we stated last class. Note that most of these follow from the additivity property 1.4. Have $A, B \subseteq S$ be events.

Theorem 2.1. If A^C is the complement of A, then $P(A^C) = 1 - P(A)$.

Proof. A and A^C are by definition disjoint, and their union is S. By additivity 1.4 $P(A) + P(A^C) = P(S) = 1$. □

Theorem 2.2. $P(A) = P(A \cap B) + P(A \cap B^C)$

The set of $\{x \in A : x \in B\}$ and $\{x \in A : x \notin B\}$ are by definition disjoint, and the union of the two is A. This then follows by additivity 1.4.

Theorem 2.3. If A contains B, $P(A) = P(B) + P(A \cap B^C)$

Proof. Have 2.2, except $P(A \cup B) = P(B)$ where $A \supseteq B$. □

Theorem 2.4 (Monotonicity)**.** If $A \supseteq B$, then $P(A) \geq P(B)$

Immediately follows from 2.3, since $P(A \cap B^C)$ must be non-negative, giving the inequality.

Theorem 2.5

Suppose $A_1, A_2, \ldots$ are a sequence of events which form a *partition* of S (pairwise disjoint) , with their union being the entire sample space ($\bigcup_i A_i = S$). Let B be any event. Then we have

$$P(B) = \sum_i P(A_i \cap B)$$

Theorem 2.6 (Principle of Inclusion-Exclusion)**.** $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Proof. The events $A \cap B^C$, $B \cap A^C$, $A \cap B$ are disjoint events.

$$\begin{aligned} P(A \cup B) &= P(A \cap B) + P(A \cap B^C) + P(B \cap A^C) \\ &= P(A \cap B) + [P(A) - P(A \cap B)] + [P(B) - P(A \cap B)] \qquad \text{from } 2.2 \end{aligned}$$

□

For a more generalized version of inclusion-exclusion formula, look at Challenge 1.3.10 in textbook.

Theorem 2.7 (Subadditivity)**.** For any sequence of events $A_1, A_2, \ldots$ not necessarily pairwise disjoint, have

$$P(A_1 \cup A_2 \ldots) \leq P(A_1) + P(A_2) + \ldots$$

> **Remark 2.8.** This is more of a worry in grad-level courses, where you study more pathological probability spaces, but 'uncountable' subadditivity does not exist. Consider $S = \text{Uniform}([0,1])$. Have $A_x = \{x\}$ for $x \in S$. $P(\bigcup_{x \in S}) A_x = P(S) = P([0,1]) = 1$. Yet for any 'singleton' x, $P(A_x) = P(\{x\}) = 0$, meaning $\sum_{x \in S} P(A_x) = 0$.

§2.2 Uniform Probabilities on Finite Spaces

Have $S = \{s_1, \ldots, s_n\}$. For all $\{s_i\}$ to have the same probability, $P(\{s_i\}) = \frac{1}{n}$, called a *discrete uniform distribution*.

Any $A \subseteq S$ with k elements, would have $P(A) = \frac{k}{n}$, meaning

$$P(A) = \frac{|A|}{|S|}$$

A problem solving technique to find P of a rather complicated event is to see if the probability of its complement can be easily found, then use 2.1.