

Lecture #2: Foundations of Probability

MATH 321 – Mathematical Statistics I

Torin Quinlivan

Knox College
tquinlivan@knox.edu

October 17, 2026



What is Probability?

A **probability** p of an event is a number $0 \leq p \leq 1$.

The **frequentist interpretation** is that p is the proportion of times the event will occur if I repeat the experiment over and over.

The **Bayesian interpretation** is that p is my degree of belief that the event will occur, given what I believe to be true. (We can update this belief as we get more information.)

What is Probability?

1. Toss a coin. What is the probability of getting a heads?
2. Throw a 6-sided die. What is the probability of getting a “1”?
3. Draw five cards from a deck. What is the probability of drawing two aces?

What is Probability?

1. Toss a coin. What is the probability of getting a heads?
2. Throw a 6-sided die. What is the probability of getting a “1”?
3. Draw five cards from a deck. What is the probability of drawing two aces?

Careful with the wording! Does this question mean:

1. What is the probability of drawing *exactly* two aces?
2. What is the probability of drawing *at least* two aces?

Probability: A Formal Approach

Let's formalize our definition of probability, using set theory:

A **probability space** is a pair (Ω, P) , where Ω is a set called the **sample space** whose elements are the **outcomes** of the experiment. Subsets $A \subseteq \Omega$ are called **events**, and $P : \mathcal{P}(\Omega) \rightarrow [0, 1]$ is a function such that

1. $\forall A \in \mathcal{P}(\Omega) (P(A) \geq 0)$, (For all events $A \subseteq \Omega$, the probability of the event A is $P(A) \geq 0$.)
2. $P(\Omega) = 1$, (The probability of the sample space Ω is 1, as some outcome must occur each experiment.)
3. $\forall A_1, A_2, A_3, \dots \in \mathcal{P}(\Omega) (A_i \text{ pairwise disjoint} \rightarrow P(\bigcup A_i) = \sum P(A_i))$ (For pairwise disjoint (or mutually exclusive) events, the probability of the union of events is equal to the sum of the probabilities of the events.)

Complement Rule

Corollary (Complement Rule):

$$\forall A \in \mathcal{P}(\Omega) \ (P(\overline{A}) = 1 - P(A)).$$

What does this mean in English?

Complement Rule

Corollary (Complement Rule):

$$\forall A \in \mathcal{P}(\Omega) \ (P(\overline{A}) = 1 - P(A)).$$

What does this mean in English?

For all events A , the probability of the complement of the event, $\overline{A}$, is 1 minus the probability of the event A .

Complement Rule

Corollary (Complement Rule):

$$\forall A \in \mathcal{P}(\Omega) \ (P(\overline{A}) = 1 - P(A)).$$

What does this mean in English?

For all events A , the probability of the complement of the event, $\overline{A}$, is 1 minus the probability of the event A .

This makes intuitive sense. If the Bears have a 46% chance of winning on Sunday, they have a $100\% - 46\% = 54\%$ chance of not winning.

Inclusion/Exclusion Principle

Corollary (Inclusion/Exclusion Principle):

$$\forall A, B \in \mathcal{P}(\Omega) \ (P(A \cup B) = P(A) + P(B) - P(A \cap B)).$$

What does this mean in English?

Inclusion/Exclusion Principle

Corollary (Inclusion/Exclusion Principle):

$$\forall A, B \in \mathcal{P}(\Omega) \ (P(A \cup B) = P(A) + P(B) - P(A \cap B)).$$

What does this mean in English?

For any two events A and B , the probability of the union of the two events is equal to the probability of A plus the probability of B minus the probability of the intersection of A and B (because it's being double counted).

Inclusion/Exclusion Principle

Corollary (Inclusion/Exclusion Principle):

$$\forall A, B \in \mathcal{P}(\Omega) \quad (P(A \cup B) = P(A) + P(B) - P(A \cap B)).$$

This can be extended to more than two events, with the pluses and minuses alternating:

$$\begin{aligned} P\left(\bigcup A_i\right) &= P(A_1) + P(A_2) + \cdots + P(A_n) \\ &\quad - P(A_1 \cap A_2) - P(A_1 \cap A_3) - \cdots - P(A_{n-1} \cap A_n) \\ &\quad + P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_4) + \cdots \\ &\quad + P(A_{n-2} \cap A_{n-1} \cap A_n) - P(A_1 \cap A_2 \cap A_3 \cap A_4) \\ &\quad \vdots \end{aligned}$$

Examples

Example: Consider rolling a fair six-sided die.

In our formal mathematical notation, let $\Omega = \{1, 2, 3, 4, 5, 6\}$ and for $x \in \Omega$, $P(x) = \frac{1}{6}$.

1. What is the probability of rolling a 1?
2. What is the probability of *not* rolling a 1?
3. What is the probability of rolling an even number?
4. What is the probability of rolling an odd number?

Equally Likely Outcomes

If $|\Omega|$ is finite and all **atomic** outcomes in Ω are *equally likely*, then for $x \in \Omega$ and $A \subseteq \Omega$:

$$P(x) = \frac{1}{|\Omega|} \quad \text{and} \quad P(A) = \frac{|A|}{|\Omega|}.$$

The probability of A is the number of outcomes in A divided by the total number of possible outcomes.

A **uniform discrete probability space** has N outcomes, each of which occurs with probability $\frac{1}{N}$.

Lots of probability spaces are uniform and you can evaluate their probabilities using the probability axioms and counting formulas.

Equally Likely Outcomes Examples

Example: Consider a standard deck of 52 cards and suppose a hand of five cards is drawn.

1. What is the probability of drawing a hand with two aces?
2. What is the probability of drawing a pair of aces and three kings?
3. What is the probability of drawing a full house?

Probability Distributions

Random Variables

A **random variable** X on Ω is a function $X : \Omega \rightarrow \mathbb{R}$. Since each $x \in \Omega$ has a probability associated with it, each $X(x)$ has a probability as well:
$$P(X = a) = P(\{x \in \Omega : X(x) = a\}).$$

Probability Mass Function

The **probability mass function** of a *discrete* random variable gives the probability that it takes on each value in its range: $f(x) = P(X = x)$.

Probability Distribution Example

Suppose we roll two dice and take the sum of the values. The probability distribution for X is:

x	2	3	4	5	6	7	8	9	10	11	12
$P(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

1. What is the probability the sum is greater than 9?
2. What is the probability the sum is odd?

Probability Distributions

Distribution functions give a way to compute probability when outcomes aren't equally likely.

Distribution functions come in two flavors: **probability mass (or density) functions** and **cumulative distribution functions**, which represents $P(X \leq x)$.

Cumulative distribution functions are most often used when dealing with continuous random variables (e.g., the normal distribution functions) because probabilities can be evaluated with integrals.

Probability distributions give us a way to classify families of similar random variables. The study of families of probability distributions occupies a large part of probability theory (and statistics).

The Binomial Distribution

Suppose an experiment that has two possible outcomes (success or failure); is (independently) repeated n times (trials); and that the probability of success (each time) is p . Let X be a random variable which counts the number of successes that occur among these n trials. Then the probability mass function is:

$$f(x) = P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x = 0, 1, 2, \dots, n$$

Example: Suppose a loaded coin which has probability of heads $p = .7$ is tossed 10 times.

1. What is the probability of getting exactly 8 heads?
2. What is the probability of getting at least 8 heads?