On Probability Theory and Stochastic Processes

September 29, 2025

Timothy Tarter
James Madison University
Department of Mathematics

Abstract

This lecture plan is intended for a four-day crash course in probability theory and stochastic processes.

Combinatorics

Consider the following questions:

Example 1. *If there are ten seats on a bus and ten students looking to board it, how many ways could they be ordered for a seating chart?*

Solution:

$$10! = 10 * 9 * 8 * 7 * 6 * 5 * 4 * 3 * 2 * 1 \tag{1}$$

Example 2. *If a locker code requires a 3-digit code using digits 0-9 where digits cannot be repeated, how many possible codes could one make?*

Solution:

$$10 \text{ permute } 3 = \frac{10!}{(10-3)!} = 10 * 9 * 8 \tag{2}$$

This leads to the permutation formula, for the case when order DOES matter.

Definition 1. *If* ***order matters*** *in picking groupings of r at a time from n things, then*

$$n \textit{ permute } r = \frac{n!}{(n-r)!} \tag{3}$$

Example 3. *A math committee has 12 members, and they want to pick 3 members to be the council for the group. Since order doesn't matter, how many ways can they pick a council?*

Solution:

$$\binom{12}{3} = \frac{12!}{(12-3)!3!} = 220 \tag{4}$$

This leads to the combination formula for when order DOES NOT matter.

Definition 2. *If **order does not matter** in picking groupings of r at a time from n things, then,*

$$\binom{n}{r} = \frac{n!}{(n-r)!r!} \tag{5}$$

Group Work

Problem 1. A password consists of 4 distinct letters chosen from the 26 English letters. How many possible passwords are there? (Order matters.)

Solution.

$$P(26,4) = \frac{26!}{(26-4)!} = \frac{26!}{22!} = 26 \cdot 25 \cdot 24 \cdot 23 = 358{,}800$$

Problem 2. A committee of 5 students is to be chosen from a class of 20. How many different committees can be formed? (Order does not matter.)

Solution.

$$\binom{20}{5} = \frac{20!}{5! \cdot 15!} = 15{,}504$$

Problem 3. There are 10 runners in a race. In how many ways can the first, second, and third place medals be awarded?

Solution.

$$P(10,3) = \frac{10!}{(10-3)!} = \frac{10!}{7!} = 10 \cdot 9 \cdot 8 = 720$$

Problem 4. A bag contains 12 marbles. You randomly pick 4 marbles. How many different groups of marbles could you pick? (Order does not matter.)

Solution.

$$\binom{12}{4} = \frac{12!}{4! \cdot 8!} = 495$$

Problem 5. Seven friends line up to take a photo. How many different arrangements of the seven friends are possible?

Solution.

$$7! = 5040$$

Bridging the Gap: Combinatorics and Probability

Definition 3. *We say that, given a sample space Ω containing all possible outcomes / events, the probability of a single event $A \subseteq \Omega$ occurring is*

$$p(x = A) = \frac{\mu(A)}{\mu(\Omega)} \tag{6}$$

where $\mu(A)$ is the measure of A (how many ways it can happen in Ω) and $\mu(\Omega)$ is the number of things in the sample space. In short,

$$p(x = A) = \frac{\textit{number of ways we can construct an event}}{\textit{number of constructions in the sample space}}. \tag{7}$$

We will delve more deeply into measure theory in two weeks; suffice it to say that a measure is a way of counting how many things are in a set. For discrete sets, this is typically combinatorial in nature, however, for continuous sets, we require a more rigorous theory of integration.

Definition 4. *Let $X : \Omega \to \mathbb{R}$ be a random variable. Then for each element x of the sample space Ω, $X(x) = v \in \Omega$, or the value assigned to x by X.*

Example 4. *Consider a fair six sided die. $\Omega = \{1, 2, 3, 4, 5, 6\}$ and X being a random variable such that*

- $P(X = 1) = \frac{1}{6}$
- $P(X = 2) = \frac{1}{6}$
- $P(X = 3) = \frac{1}{6}$
- $P(X = 4) = \frac{1}{6}$
- $P(X = 5) = \frac{1}{6}$
- $P(X = 6) = \frac{1}{6}$.

We call this list the "probability distribution of X". Notably, all probability distributions must sum (or integrate) to one.

Definition 5. *A probability distribution p for a random variable X over the sample space Ω has the property that*

1. $p(X = k) > 0$ *for any* $k \in \Omega$
2. $\sum_{k \in \Omega} p(X = k) = 1$

Definition 6. *The expected value of a probability distribution is defined by*

$$E[X] = \sum_{k \in \Omega} p(X = k) * k. \tag{8}$$

This is sometimes called the "first moment" of the distribution.

Example 5. *For example 4, we can see that*

$$E[X] = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = 3.5. \tag{9}$$

Definition 7. *The variance of a distribution (or the "second moment" of the distribution) is defined by*

$$V[X] = E[X^2] - E[X]^2. \tag{10}$$

This leads to an important note, we can calculate $E[X^2]$ as follows:

$$E[X^2] = \sum_{k\in\Omega} p(X = k) * k^2. \tag{11}$$

Example 6. *For example 4,*

- $E[X]^2 = 3.5^2 = 12.25$
- $E[X^2] = \frac{1}{6}(1 + 4 + 9 + 16 + 25 + 36) = 15.1\overline{66}$

thus,

$$V[X] = 15.1\overline{66} - 12.25 = 2.91\overline{66}. \tag{12}$$

We will see more on this tomorrow with continuous distributions, but $E[X]$ is what we traditionally think of as the "mean" and $\sqrt{V[X]} = \sigma$ is the conventional notion of "standard deviation." This will eventually lead us to ask the question... can we approximate discrete distributions with the normal distribution? (Short answer, yes!).

Discrete Probability Distributions

The Uniform Distribution

We originally introduced the idea of the uniform distribution with the fair six-sided die. If X is a uniformly distributed random variable, then every element of Ω is equally likely to be the value of X. I.e.,

$$P(X = k) = \frac{1}{\mu(\Omega)}. \tag{13}$$

The Binomial Distribution

Definition: The Binomial distribution models the number of successes in n independent Bernoulli trials, each with probability of success p.

Probability Mass Function (PMF):

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \ldots, n$$

Expectation and Variance:

$$E[X] = np, \quad \text{Var}(X) = np(1-p)$$

The Geometric Distribution

Definition: The Geometric distribution models the number of trials until the first success, with independent Bernoulli trials each having probability of success p.

Probability Mass Function (PMF):

$$P(X = k) = (1 - p)^{k-1} p, \quad k = 1, 2, 3, \dots$$

Expectation and Variance:

$$E[X] = \frac{1}{p}, \quad \text{Var}(X) = \frac{1 - p}{p^2}$$

The Hypergeometric Distribution

Definition: The Hypergeometric distribution models the number of successes in n draws, without replacement, from a finite population of size N containing K successes.

Probability Mass Function (PMF):

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}, \quad \max(0, n - (N - K)) \leq k \leq \min(n, K)$$

Expectation and Variance:

$$E[X] = n \cdot \frac{K}{N}, \quad \text{Var}(X) = n \cdot \frac{K}{N} \cdot \left(1 - \frac{K}{N}\right) \cdot \frac{N - n}{N - 1}$$

The Negative Binomial Distribution

Definition: The Negative Binomial distribution models the number of trials needed to achieve r successes, with independent Bernoulli trials each having probability of success p.

Probability Mass Function (PMF):

$$P(X = k) = \binom{k - 1}{r - 1} p^r (1 - p)^{k-r}, \quad k = r, r + 1, r + 2, \dots$$

where X is the number of trials required to obtain r successes.

Expectation and Variance:

$$E[X] = \frac{r}{p}, \quad \text{Var}(X) = \frac{r(1 - p)}{p^2}$$

The Poisson Distribution

Definition: The Poisson distribution models the number of events occurring in a fixed interval of time or space, when events occur independently and at a constant average rate $\lambda > 0$.

Probability Density Function (PMF):

$$f(k;\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

Expectation and Variance:

$$E[X] = \lambda, \quad \text{Var}(X) = \lambda$$

The Big Four Table:

Distribution	**Scenario**	**Key Assumptions**	**Typical Use Cases**
Binomial	Fixed number of n independent trials with success probability p	Each trial is independent; probability of success is constant p	Number of defective items in a batch, number of heads in n coin flips, number of “yes" votes in a poll
Negative Binomial	Count trials until r successes occur	Each trial is independent; probability of success is constant p	Number of tosses of a coin needed to get r heads, number of patients until r recoveries are observed in a clinical trial
Geometric	Special case of Negative Binomial with $r = 1$ (count trials until the *first* success)	Independent trials; probability of success p constant	Number of die rolls until the first six appears, number of phone calls until first answered
Hypergeometric	Draw n items without replacement from a finite population of size N with K successes	Sampling without replacement (dependence between draws)	Drawing colored balls from an urn, quality control sampling (e.g., number of defectives in a sample from a finite lot)

Table 1: Comparison of Use Cases for Discrete Distributions

Group Work

1. A fair coin is flipped 10 times. What is the probability of getting exactly 6 heads?

2. In a basketball game, a player makes a free throw with probability 0.75. Out of 8 attempts, what is the probability she makes at least 6 shots?

3. Suppose you roll a fair die repeatedly until a six appears. What is the probability that the first six comes on the fourth roll?

4. A computer server crashes with probability 0.02 on any given day, independently of other days. What is the expected number of days until the first crash?

5. A baseball player has a batting average of 0.3. What is the probability that he gets his third hit on the seventh at-bat?

6. A card player draws cards with replacement, where the probability of drawing a heart is 0.25. What is the expected number of draws until she obtains 5 hearts?

7. A lot contains 100 items, of which 20 are defective. If 10 items are sampled without replacement, what is the probability that exactly 3 are defective?

8. A box has 12 red balls and 8 blue balls. If 5 are drawn without replacement, what is the probability of getting all red?

9. Which distribution (Binomial, Geometric, Negative Binomial, or Hypergeometric) would be appropriate in each case? (a) Rolling a die n times and counting the number of 6's. (b) Counting the number of die rolls until the first 6. (c) Drawing n cards from a deck without replacement and counting the number of hearts. (d) Counting the number of die rolls until the 5th 6.

10. A quality inspector samples 4 lightbulbs at random from a batch of 50, which contains 5 defectives. What is the probability that at least one defective is found?

Solutions Guide

1. Binomial with $n = 10$, $p = 0.5$:

$$P(X = 6) = \binom{10}{6}(0.5)^6(0.5)^4 = \frac{210}{1024} \approx 0.205.$$

2. Binomial with $n = 8$, $p = 0.75$:

$$P(X \geq 6) = \sum_{k=6}^{8} \binom{8}{k}(0.75)^k(0.25)^{8-k}.$$

   Numerical value ≈ 0.594.

3. Geometric with $p = \frac{1}{6}$:

$$P(X = 4) = (5/6)^3 \cdot (1/6) \approx 0.0965.$$

4. Geometric expectation:

$$\mathbb{E}[X] = \frac{1}{p} = \frac{1}{0.02} = 50 \text{ days.}$$

5. Negative Binomial with $r = 3$, $p = 0.3$, $k = 7$:

$$P(X = 7) = \binom{6}{2}(0.3)^3(0.7)^4 \approx 0.097.$$

6. Negative Binomial expectation:

$$\mathbb{E}[X] = \frac{r}{p} = \frac{5}{0.25} = 20 \text{ draws.}$$

7. Hypergeometric with $N = 100$, $K = 20$, $n = 10$, $k = 3$:

$$P(X = 3) = \frac{\binom{20}{3}\binom{80}{7}}{\binom{100}{10}}.$$

   Approximate value ≈ 0.238.

8. Hypergeometric with $N = 20$, $K = 12$, $n = 5$, $k = 5$:

$$P(X = 5) = \frac{\binom{12}{5}\binom{8}{0}}{\binom{20}{5}} \approx 0.0165.$$

9. (a) Binomial, (b) Geometric, (c) Hypergeometric, (d) Negative Binomial.

10. Hypergeometric with $N = 50$, $K = 5$, $n = 4$:

$$P(\text{at least one defective}) = 1 - P(0 \text{ defective}) = 1 - \frac{\binom{45}{4}}{\binom{50}{4}} \approx 0.339.$$