Chapter 1

Introduction: Five Questions

Lecture 1 | Devore 9e: Ch. 1 (overview)

The whole course is previewed by five questions. None of them can be answered today. Each one names the tool that will answer it, and the lecture where that tool arrives.

1.1 Question 1: Poker

Five cards are dealt from a standard 52-card deck. The hands, from best to worst, begin

$$\text{royal flush} > \text{straight flush} > \text{four of a kind} > \text{full house} > \text{flush} > \text{straight} > \cdots$$

Why is that the ranking? Why is a flush better than a straight?

The ranking is not a convention — it is a *fact*, forced by how rare each hand is. To see that, we must be able to count how many five-card hands of each type exist. Counting well is harder than it looks.

→ **Lecture 3** (counting and combinatorics), where we compute all of these.

1.2 Question 2: How Buffon computed π

Rule a floor with parallel lines a distance a apart, and drop a needle of length $\ell < a$ on it N times. Let n be the number of times the needle crosses a line. Then — as we will prove —

$$\frac{n}{N} \approx \frac{2\ell}{\pi a} \qquad \Longrightarrow \qquad \pi \approx \frac{2\ell N}{a\, n}.$$

You can compute π by dropping a needle on the floor. Why does that work?

Note that the sample space here is not a finite list of outcomes — the needle can land in a continuum of positions and angles. Counting is useless; we need to measure *area* instead.

→ **Lecture 4** (geometric probability), where this is derived.

1.3 Question 3: The flu test

A test for the flu is

- correct 99% of the time on people who *have* the flu;
- correct 95% of the time on people who *do not*.

You take the test and it comes back positive.

What is the probability that you have the flu?

Most people answer "about 99%". That answer is not just wrong, it is not even well posed: **the question cannot be answered from the information given.** You also need to know how common the flu is right now. If the flu is rare, a positive result may well mean less than a coin flip.
→ **Lectures 5–7** (conditional probability and Bayes' theorem).

1.4 Question 4: Two ways to summarize an experiment

(a) Toss a coin three times. The outcomes are $HHH, HHT, HTH, HTT, THH, THT, TTH, TTT$ — eight of them. But usually we do not care *which* sequence occurred, only how many heads it contained. Let X be that number. Then X takes the values $0, 1, 2, 3$ with probabilities

$$\tfrac{1}{8}, \quad \tfrac{3}{8}, \quad \tfrac{3}{8}, \quad \tfrac{1}{8}.$$

Eight outcomes have been compressed into four numbers.

(b) Wait for a bus. Let X be the number of minutes you wait. Now X can be 3, or 3.5, or $3.14159\ldots$ — any real number in an interval. The probability of waiting *exactly* 3 minutes, to infinite precision, is zero. So the list-the-values-and-their-probabilities approach breaks down completely.

What is the right way to describe a numerical quantity that is random?

→ **Lectures 7–8** for case (a), **Lecture 13** for case (b).

1.5 Question 5: What score do you need?

Suppose the STAT400 final has mean 88 and variance 15.

You want to finish in the top 10% *of the class. What score should you aim for?*

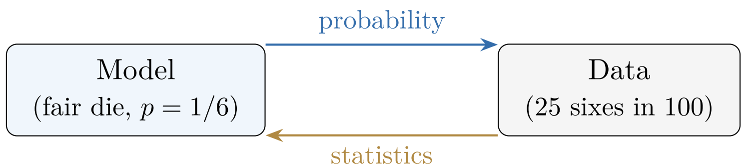
The mean and the variance alone are not enough — you also need to know the *shape* of the distribution of scores. Once we know that shape is the normal curve, the answer is a short computation:

$$\sigma = \sqrt{15} \approx 3.87, \qquad x = \mu + z_{0.90}\,\sigma \approx 88 + (1.28)(3.87) \approx 93.$$

→ **Lecture 14** (the normal distribution).

1.6 The shape of the course

Two of those five questions (1 and 2) are about *probability*: the mechanism is known, and we ask what the data will look like. The last part of the course reverses the arrow — the data is observed and the mechanism is unknown.



Lectures	**Block**	**Content**
2–7	Probability	sets, counting, axioms, conditional probability, independence, Bayes
8–12	Random variables	pmf, cdf, expectation, variance; the named discrete families
13–17	Continuous rv's	pdf, transformations; normal, exponential, gamma, Weibull, beta
18–21	Joint behavior	joint distributions, covariance, correlation, CLT
22–26	Statistics	estimators, bias, method of moments, maximum likelihood

Remark 1.1 (Notation)**.** S is the sample space, ω a single outcome, A, B, E events, $\mathbb{P}$ the probability. Capital letters near the end of the alphabet (X, Y, Z) are random variables; the matching lower case (x, y, z) are the values they take. In lecture I often write ξ (xi) for a random variable; it means the same as X.