Probability and Statistical Inference

Tianqi Zhang
Emory University

Apr 17th 2025

2 Sampling

2.1 Sample Spaces and Probability Models

To describe randomness formally, we need a language that specifies all the possible outcomes and assigns probabilities to them. This is the role of a probability model.

Definition 2.1.1 (Sample Space). A **sample space** Ω is the set of all possible outcomes of a random experiment. A probability model also requires a rule that assigns probabilities to events, which are subsets of Ω.

Let's look at a few examples to make this concrete.

Example 2.1.2 Coin Tossing In a single coin toss, the outcome is either heads (H) or tails (T). We write:

$$\Omega = \{\mathrm{H}, \mathrm{T}\}.$$

If the coin is fair, it makes sense to assign:

$$P(\mathrm{H}) = P(\mathrm{T}) = \frac{1}{2}.$$

We can extend this assignment to events involving multiple outcomes. For example,

$$P(\{\mathrm{H} \text{ or } \mathrm{T}\}) = 1$$

This kind of modeling is supported by empirical data. By the **Law of Large Numbers (LLN)**, if we toss the coin many times and track how often heads appears, we should see the relative frequency stabilize around 0.5:

$$P(A) \approx \lim_{n\to\infty} f_n(A).$$

We'll revisit this idea — and formally define the limit plim — later.

2.2 Uniform Probability Models and Discrete Distributions

The coin toss is an example of a uniform probability model: each outcome is equally likely. Let's generalize this to any finite or countable space.

Example 2.2.1 Uniform Distribution Suppose the sample space is:

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}.$$

Then a uniform probability model assigns equal weight to each outcome:

$$P(\omega_i) = \frac{1}{N}, \quad \text{for all } i = 1, \ldots, N.$$

More generally, for any event $A \subseteq \Omega$, we define:

$$P(A) = \frac{|A|}{|\Omega|}.$$

This approach works well for simple random experiments like rolling a die. For example:

$$\Omega = \{1, 2, 3, 4, 5, 6\}, \quad P(i) = \frac{1}{6}.$$

Let A be the event that the outcome is even. Then:

$$A = \{2, 4, 6\}, \quad P(A) = \frac{3}{6} = \frac{1}{2}.$$

Remark 2.1 *On Notation Even though we write $P(\omega_i)$ for convenience, we are technically assigning probability to the set $\{\omega_i\}$, not the element itself. That is,*

$$P(\omega_i) \equiv P(\{\omega_i\}),$$

since events are subsets of Ω.

2.3 Why the Sample Space Matters

In some problems, how we define Ω completely changes the model.

Example 2.3.1 The Twin Paradox Suppose a family has two children, and we're told that one of them is a girl. What is the probability that both are girls?

Let's break it into two interpretations:

Case 1: We are told the gender of the first child. The second child is still random:

$$\Omega = \{\mathrm{G}, \mathrm{B}\}, \quad P(\mathrm{G}) = \frac{1}{2}.$$

Case 2: We are told at least one child is a girl, but not which one. Then the possible combinations are:

$$\Omega = \{\mathrm{GG}, \mathrm{GB}, \mathrm{BG}\}.$$

Each outcome is equally likely, so:

$$P(\mathrm{GG}) = \frac{1}{3}.$$

This example illustrates how the structure of Ω — and what we know — affects the probability model.

2.4 Bernoulli Trials and the Binomial Distribution

Let's now consider what happens when we repeat a simple random experiment multiple times — like flipping a coin, answering a yes/no question, or testing whether a lightbulb works.

Definition 2.4.1 (Bernoulli Trial). A **Bernoulli trial** is a random experiment with only two possible outcomes, usually called "success" (1) and "failure" (0). For example, in a coin flip:

$$\Omega = \{0, 1\}, \quad P(1) = p, \quad P(0) = 1 - p.$$

Here, p is the probability of success.

Now suppose we repeat the experiment n times independently. The full sample space is then:

$$\Omega^n = \{0, 1\}^n,$$

which consists of all sequences of 0s and 1s of length n. For example, a possible outcome could be $(1, 0, 1, 1, 0)$ — meaning 3 successes and 2 failures.

We define a function that counts the number of successes: Let $S_n : \Omega^n \to \{0, 1, \ldots, n\}$ be the function

$$S_n(\omega) = \sum_{i=1}^{n} \omega_i.$$

This counts how many of the n trials resulted in success.

What is the probability that we get exactly x successes in n trials?

Each specific sequence with x successes and $(n-x)$ failures has probability:

$$p^x(1-p)^{n-x}.$$

And how many such sequences are there? That's given by a binomial coefficient:

Definition 2.4.2 (Binomial Coefficient).

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

is the number of different sequences (orderings) that contain exactly x successes in n trials.

Putting this together, we get the **binomial distribution**:

Definition 2.4.3 (Binomial Distribution). The probability of observing exactly x successes in n independent Bernoulli trials, each with success probability p, is:

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

This is called the **binomial distribution** with parameters n and p, written:

$$X \sim Binom(n, p).$$

Remark 2.2 *The binomial formula accounts for both the probability of each individual sequence and the number of such sequences. The sum of all probabilities over $x = 0$ to n equals 1.*