# Probability and Statistical Inference

Tianqi Zhang
Emory University

Apr 17th 2025

## 2 Sampling

### 2.1 Sample Space and Probability Models

**Definition 2.1.1 (Sample Space).**
The specification of a probability model requires:

1. A sample space, $\Omega$: The set of all possible outcomes in the problem.
2. A probability assignment for these outcomes (subsets of $\Omega$).

Consider the following examples to help understanding the construction of probability models:

**Example 2.1.2 Coin Tossing**

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

- If assuming the coin is balanced, we state $P(\mathrm{H}) = P(\mathrm{T}) = \frac{1}{2}$.
- This assumption can be extended to all subsets of $\Omega$.
- Assumption can be justified empirically using the **Law of Large Numbers (LLN)**:

$$P(A) = \operatorname*{plim}_{n\to\infty} f_n(A).$$

For a proper definition of "plim"

To generalize the above results:

**Definition 2.1.3 (Uniform probability measure/distribution).**

- $\Omega = \{\omega_1, \dots, \omega_N\}$ or countable
- **Uniform probability measure (distribution)**: For any $A \subseteq \Omega$,

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

  and define

$$P(\omega_i) = \frac{1}{N}, \forall i \in \{1, 2, \dots, N\}$$

- **Warning on abuse of notation:** Note that this is not a proper probability measure, as it assigns probabilities to individual elements, not subsets.

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

  even though $\omega_i \in \Omega$, $\{\omega_i\} \subseteq \Omega$, and $\{\omega_i\} \in \mathscr{P}(\Omega)$ as the power set.

Therefore, we have the above function $P$ defined to be $\mathscr{P}(\Omega) \to [0, 1]$

**Example 2.1.4 Fair Die**

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

- Let $A =$ "even numbers" $= \{2, 4, 6\}$. Then:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{3}{6} = \frac{1}{2}.$$

**Remark 2.1** *Conditioning and Specifying $\Omega$ is crucial*

**Example 2.1.5 Twin paradox example**

- **Assumptions:**
  - (i) Gender of newborn: $P(\text{Girl}) = P(\text{Boy}) = \frac{1}{2}$.
  - (ii) Gender of one child is independent of the gender of the other child.
- Known: A family has two children, one of whom is a girl.
- Question: What is $P(\text{both children are girls})$?

- Case (i): If we know the gender of the first child, then $\Omega$ considers only the second child

$$\Omega_i = \{\text{G}, \text{B}\}.$$

Then:

$$P(\text{G}) = \frac{1}{2}.$$

- Case (ii): If we know at least one child is a girl:

$$\Omega_{ii} = \{\text{GG}, \text{GB}, \text{BG}\}.$$

Then:

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

## 2.2 Bernoulli Trials and Binomial Distribution

**Example 2.2.1 Bernoulli Distribution**

- Suppose that we have a sequence of $n$ independent experiments

$$P(\text{Success}) = p,\ P(\text{Failure}) = 1 - p$$

- We have the sample space:

$$\Omega = \{0, 1\}^n \quad \text{(Cartesian product of the simple } \{0, 1\}\text{)}.$$

- A typical element (draw) $\omega \in \Omega$ is $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ as a sequence.
- Define $S_n : \Omega \to \mathbb{R}$, $S_n(\omega) = \sum_{i=1}^{n} \omega_i$ to be the count of success.

**Claim 2.2.2:**

Then the **independent experiment** with its probability measure:

$$P(\omega) = p^{S_n(\omega)}(1 - p)^{n - S_n(\omega)}$$

**Remark 2.2** *$S_n$ is a **random variable**, a function from $\Omega$ endowed with the probability measure $P$ and maps outcomes to $\mathcal{S} \equiv \{0, 1, \ldots, n\}$.*

Since the original Bernoulli trial is simple enough to be binary: $\Omega = \{0, 1\}$, the count of success can be reduced into the simple sum in the sequence. Therefore, we can then define a new probability measure (distribution) $P^{S_n}$ on $S$ such that:

**Theorem 2.2.3 Push-Forward Measure**
We can then define a new probability measure (distribution) $P^{S_n}$ on $S$ such that

$$\forall x \in S, P(x) \equiv \sum_{\omega \in S_n^{-1}(x)} P(\omega) = P\left[S_n^{-1}(x)\right] \equiv P^{S_n}(x)$$

Where

$$S_n^{-1}(A) \equiv \{\omega \in \Omega, S_n(\omega) \in A\} \quad \text{is the pre-image}$$

For computation purposes:

$$\forall x \in \mathcal{S},\ P^{S_n}(x) = \left|S_n^{-1}(x)\right| p^x (1-p)^{n-x}$$

Intuitively speaking, the pre-image $S_n^{-1}(x)$ indicates all possible sequences $\omega$ such that they contain exactly $x$ counts of success. We give the calculation as follows

**Definition 2.2.4 (Binomial number).**

$$\left|S_n^{-1}(x)\right| = \binom{n}{x} = \frac{n!}{x!(n-x)!} = \binom{n}{n-x}$$

To be the number of ways to pick $x$ elements from $n$ elements without order.

**Definition 2.2.5 (Binomial Distribution).**
Binomial Distribution: $Binom(n, p)$ is the push-forward probability measure of the Bernoulli trial with $n$ experiments and success probability $p$ via the random variable $S_n$ which counts success.

$$\forall x \in \mathcal{S}, P^{S_n}(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

and

$$\sum_{n=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} = 1$$

We say the **random variable** $S_n$ **follows the binomial distribution**, $S_n \sim Binom(n, p)$.