Introduction to Probability and Statistics (80430)

Ohad Feldheim

Moshe Krumbein

Fall 2022

Compiled: 2023-03-17 12:46:50

https://github.com/outofink/notes

CONTENTS

1 Introduction **1**

1.1 What is Statistics? 1

1.2 Trials 1

1.2.1 Examples 1

1.3 Characteristics of the Probability Function 2

1.4 Examples 2

1.5 Two-step Experiments 4

1.6 The Birthday Paradox 5

1.7 Law of Total Probability 5

INTRODUCTION TO PROBABILITY AND STATISTICS

1

INTRODUCTION

Moshe Krumbein - Fall 2022

1.1 What is Statistics?

The quantitative measure of the amount of uncertainty in regards to an outcome within a well defined space. (How “surprised” I would be at a given outcome)

By doing a test a large number of times and counting the number of “random” outcomes, we are able to predict with certainty.

For example, we are find the boiling point of a given liquid, even though we don’t know the specific details of a given particle within the liquid.

1.2 Trials

Definition (Sample Space)**.** A *sample space* is non-empty set Ω. (That describes the results of a sampling [and potentially additional results])

1.2.1 Examples

Definition (Probability Mass Function)**.** $p : \Omega \to [0, 1]$ such that $\sum_{\omega \in \Omega} p_\omega = 1$.

$\{\omega : p_\omega > 0\}$ is called the *support* of p.

Definition (Event)**.** $\mathcal{F} = 2^\Omega$. An *event* in F is the set $E \in \mathcal{F}$

Practically, an *event* is a “subset” of Ω.

Definition (Probability Function)**.** $\mathbb{P} : \mathcal{F} \to [0, 1]$ such that $\mathbb{P}(\Omega) = 1$.

Sampling	Ω	p
Fair coin toss	$\{H, T\}$	$p_H = p_T = \frac{1}{2}$
Unfair coin toss	$\{H, T\}$	$p_H = p = 1 - p_T$
3 coin tosses	$\{H, T\}^3$	$\forall a, b, c \in \{H, T\} : p_{abc} = \frac{1}{8}$
Coin tosses until heads	$\{H, T\}^{\aleph_0}$ / $\mathbb{N} \cup \{\infty\}$	$p_i = 2^{-i}$
Picking an angle on a circle	$[0, 2\pi)$	hard
Rolling a D6 and flipping a coin	$[6] \times \{H, T\}$	$p_{dc} = \frac{1}{12}$

$$\forall (E_i)_{i\in\mathbb{N}} E_i \in \mathcal{F} : E_i \cap E_j = \emptyset \implies \mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{n=1}^{\infty} \mathbb{P}(E_i)$$

1.3 Characteristics of the Probability Function

- Additivity- $\mathbb{P}(E_1 \mathbin{\dot\cup} E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2)$
- Monotonocity - $E_1 \subset E_1 \implies \mathbb{P}(E_2) \leq \mathbb{P}(E_2)$
- Completeness - $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ (where $A^c = \Omega \setminus A$)

The trio $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is a *sample space*, $\mathcal{F}$ is a *event space* and $\mathbb{P}$ is a *probability function* on $\mathcal{F}$ is called a *probability space.*

Definition (Discrete Probability Space)**.** For p we can say that $\mathbb{P}$ fits p if $\mathbb{P}(A) = \sum_{a\in A} p_a$. If $\mathbb{P}$ has a discrete function, it will be $p_\omega = \mathbb{P}(\{\omega\})$.

Claim. The following are equivalent:

- $\mathbb{P}$ has a discrete p
- $\mathbb{P}$ is supported by a countable set
- $\mathbb{P}(A) = \sum_{a\in A} \mathbb{P}(\{a\})$

$(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{P}$ satisfies this is called a *discrete probability space.*

1.4 Examples

Example 1.4.1. In a pot there are six balls, on four of them are written A and on two are written B. One ball is removed and its letter is written down.

Description A:

$$\Omega = \{A, B\} \quad p_A = \frac{4}{6} = \frac{2}{3}, p_B = \frac{2}{6} = \frac{1}{3}$$
$$\mathbb{P}(\emptyset) = 1 - \mathbb{P}(\Omega)$$
$$\mathbb{P}(\{A\}) = 1 - \mathbb{P}(\{B\}) = \frac{2}{3}$$

Description B:

$$\Omega = \{1, \ldots, 6\} = [6] \quad \forall i \neq j : p_i = p_j \quad \forall i : p_i = \frac{1}{6}$$
$$\mathbb{P}(\text{“}A\text{”}) = \mathbb{P}([4]) = \frac{4}{6} = \frac{2}{3}$$

Definition (Bernoulli trial)**.** *Probability space* such that:

$$\left(\{0, 1\}, 2^{\{0,1\}}, \mathbb{P}_p\right)$$

where $p_0 = 1 - p, p_1 = p$.

In other words where the outcomes are either 0 or 1 and the odds of 1 is p.

Definition (Uniform Probability Space)**.**

$$(\Omega, 2^{\Omega}, \mathbb{P}_p)$$

where $\forall \omega \in \Omega : p_\omega = \frac{1}{|\Omega|}$.

A probability space is uniform *over A* if for all spaces $A \subseteq \Omega$:

$$p_a = \begin{cases} \frac{1}{|A|} & a \in A \\ 0 & a \notin A \end{cases}$$

Example 1.4.2. Sum of two D6s

Claim. Over a *uniform probability space*:

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

Definition (Multiplying Probability Spaces)**.** Let there be $\Omega_1, \Omega_2, p^1, p^2$:

$$p : \Omega_1 \times \Omega_2 \to [0, 1]$$

such that:

$$p_{(a,b)} = p^1_a \cdot p^2_b$$

Then the probability space

$$(\Omega = \Omega_1 \times \Omega_2, 2^{\Omega}, \mathbb{P}_p)$$

is the *multiplication* of:

$$(\Omega_1, 2^{\Omega_1}, \mathbb{P}_{p^1}) \times (\Omega_2, 2^{\Omega_2}, \mathbb{P}_{p^2})$$

Definition (Multiplication of Events)**.** Let $\Omega_1 \times \Omega_2$ be a multiplication of sample spaces. The event $A \times B : A \in \Omega_1,\ B \in \Omega_2$ is called a multiplication of events.

1.5 Two-step Experiments

Example 1.5.1. We flip 2 coins and we pick a number between 1 and $5h + 5$ such that h is the number of heads flipped. What are the odds the number will be 3 or less?

There are two sample spaces:

- Ω_1: number of heads
- Ω_2: number we pick

Let $(\Omega_1, \mathcal{F}, \mathbb{P})$ (the first step of the experiment), Ω_2 a sample space and for all $\omega \in \Omega_1$ the probability space $(\Omega_2, \mathcal{F}, \mathbb{P}_{2,\omega})$.

Then define $(\Omega, \mathcal{F}, \mathbb{P}) : \Omega : \Omega_1 \times \Omega_2$ where p is defined as

$$p(\omega_1, \omega_2) = \mathbb{P}_1(\{\omega_1\}) \cdot \mathbb{P}_{2,\omega_1}(\{\omega_2\})$$

In other words it's the odds of each case in the first step times the odds of the second step, given each potential first step.

The odds that the number that is picked is less than 3 is expressed by:

$$\{(\cdot, b) : b \leq 3\}$$

so we define:

$$\begin{aligned} A_0 &= \{(0, b) : b \leq 3\} \\ A_1 &= \{(1, b) : b \leq 3\} \\ A_2 &= \{(2, b) : b \leq 3\} \\ E &= A_1 \uplus A_2 \uplus A_3 \end{aligned}$$

From "additivity":

$$\mathbb{P}(E) = \mathbb{P}(A_0) + \mathbb{P}(A_1) + \mathbb{P}(A_2)$$

where:

$$\mathbb{P}(A_0) = \mathbb{P}(\{(0,b) : b \leq 3\}) = \mathbb{P}(\{(0,1),(0,2),(0,3)\}) = 3\left(\frac{1}{4}\cdot\frac{1}{5}\right)$$
$$\mathbb{P}(A_1) = 3\left(\frac{1}{2}\cdot\frac{1}{10}\right)$$
$$\mathbb{P}(A_2) = 3\left(\frac{1}{4}\cdot\frac{1}{15}\right)$$
$$\mathbb{P}(E) = 3\left(\frac{3}{60}+\frac{1}{60}+\frac{1}{60}\right) = \frac{21}{60}$$

Example 1.5.2.

$$(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$$
$$p(\omega_1, \omega_2, \omega_3, \cdots) = \mathbb{P}(\{\omega_1\})\cdot\mathbb{P}_{\omega_1}(\{\omega_2\})\cdot\mathbb{P}_{\omega_1,\omega_2}(\{\omega_2\})\cdot\cdots$$

1.6 The Birthday Paradox

$$\Omega = [265]^2$$
$$P(A^c) = \frac{|A^c|}{|\Omega|} = \frac{365\cdot 365\cdot 363\cdot(265-n+1)}{365^n}$$

n	$\mathbb{P}(A)$
23	0.507
50	0.970
70	0.999

1.7 Law of Total Probability

$$\mathbb{P}(B) = \sum_{n\in\mathbb{N}} \mathbb{P}(B\cap A_n)$$