

# MEASURE THEORY AND THE PROBABILITY OF AN INFINITE SEQUENCE OF HEADS

KENT SLINKER

revised and corrected May 2016

## I. Introduction.

The traditional theory of probability seems to have arisen purely from the desire to solve problems related to gambling. According to the mathematician Tom M. Apostol, it was a gambler's dispute in 1664 that led to the general theory of probability by way of correspondence between Blaise Pascal and Pierre de Fermat concerning a game of dice, known today as chuck-a-luck, where the question was, What was the probability of the occurrence of at least one "double six" in 24 throws of the dice?<sup>1</sup> We will not treat this problem here, but rather refer the reader to excellent treatments of the problem found in [4] and [7]. The context of Probability as applied to gambling by its very nature dealt with problems where the total number of possible outcomes, known as the sample space, was finite. Eventually questions arose which required dealing with an infinite sample space. This, together with Cantor's discovery that the cardinality of the interval  $(0, 1)$  was, in some sense, larger than the cardinality of the natural numbers, required tackling cases where the sample space was uncountably infinite. As both the applied and theoretical aspects of probability grew in complexity, and mathematicians began to worry more about foundational issues in the early parts of the 20th century, it became necessary to develop the theory in a formal and rigorous way. This was achieved largely by the work of the Russian mathematician, A. Kolmogorov, with his *Foundations of Probability Theory*, published first in Russian in 1933. Kolmogorov's approach was axiomatic in nature, and in time became part of a larger discipline known as measure theory. In this paper we will state the basic axioms of probability and examine how results from measure theory provide a proof that the probability for an infinite number of coin tosses which all come up heads is zero.

## II. The Axioms of Probability for Finite Samples Spaces.

Before we give the definition of probability for finite sample spaces, we need to define two concepts central to probability, that of a set function which is **finitely additive**, and a **sigma algebra of sets**.

**Definition** (Finitely Additive Set Function). Let  $\mathcal{F}$  be a collection of sets, then a set function  $P : \mathcal{F} \rightarrow \mathbb{R}$  is said to be **finitely additive** iff

$$P : (C \cup D) = P(C) + P(D)$$

whenever  $C$  and  $D$  are disjoint sets in  $\mathcal{F}$  such that  $C \cup D$  is also in  $\mathcal{F}$ .

---

<sup>1</sup>[1] pg. 469

Simple examples of finitely additive set functions abound. For example, let  $\mathcal{F}$  be the set of intervals in  $\mathbb{R}$  of the form  $(a, b)$  and define  $P : \mathcal{F} \rightarrow \mathbb{R}$  by  $b - a$ , then  $P$  is a finitely additive set function. The last criteria of the above definition requires closure of the sets in  $\mathcal{F}$  under the operation of taking unions, which leads to our next definition of a sigma algebra of sets.

**Definition** (Sigma algebra of sets). A non-empty class  $\mathcal{F}$  of subsets of a given universal set  $\Omega$  is called a **sigma algebra** if for every  $D$  and  $C$  in  $\mathcal{F}$  we have

$$D \cup C \in \mathcal{F} \quad \text{and} \quad D^c \in \mathcal{F}$$

Where  $D^c = \Omega \setminus D$ .

There are many consequences that can be deduced fairly easily from this above definition, for instance that  $\emptyset \in \mathcal{F}$ , (since  $\mathcal{F}$  contains at least one element  $D$ , then  $D = D \cup D \in \mathcal{F}$ , so  $D^c \in \mathcal{F}$ , hence  $\emptyset = D \cup D^c \in \mathcal{F}$ .) Note that the term sigma algebra refers to a non-empty class of subsets of a given universal set that possesses the above properties, rather than special operations defined on sets. It should be clear that if  $\mathcal{F}$  consists of all subsets of some non-empty Universal set  $S$ , then  $\mathcal{F}$  is a sigma algebra.

**Definition** (Probability for Finite Sample Spaces). Let  $\mathcal{F}$  denote a sigma algebra whose elements are subsets of a given finite set  $\Omega$ . A function  $P : \mathcal{F} \rightarrow \mathbb{R}$  is called a probability function (or measure) if it satisfies the following three conditions:

- i*)  $P$  is finitely additive
- ii*)  $P(A) \geq 0$  for all  $A \in \mathcal{F}$
- iii*)  $P(\Omega) = 1$

Given the above conditions, the triple  $(\Omega, \mathcal{F}, P)$  is called a **finite probability space**.

*Terminology.* The set  $\Omega$  is called the **sample space**, and each  $\omega \in \Omega$  is called an **outcome**, and the subsets  $A \subseteq \mathcal{F}$  are called **events**. Let  $A$  be an event, then given an outcome  $\omega$ , either  $\omega \in A$  or  $\omega \in A^c$ , in the former case we say the event  $A$  has occurred, otherwise we say the event  $A$  has not occurred. An event  $A$  is said to be impossible if  $A = \emptyset$ , since clearly  $\omega \in A^c$  for all  $\omega \in \Omega$  and certain if  $A = \Omega$ . The non-negative number,  $P(A)$ , is said to be the probability that an outcome  $\omega$  is an element of  $A$ . It is also customary to say that  $P(A)$  is the probability that the event  $A$  occurs. It is important to note that while  $P(\emptyset) = 0$ , not all events with probability 0 need be impossible.

Usually the sigma algebra is taken to be the set of all subsets of  $\Omega$ , in which case if  $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$  by properties *i*) and *ii*), we have

$$(1) \quad P(\Omega) = \sum_{i=1}^n P(\omega_i) = 1$$

Since every subset  $A \in \mathcal{F}$  can be written as some disjoint union of the  $\omega_i$ , the function  $P$  is completely determined by the value of  $P(\omega_i)$  for each  $\omega_i \in \Omega$ . Notice that one need not assign a positive probability to every element  $\omega_i \in F$ , as long as (1) holds.

Here is helpful to consider two different probability spaces with the same sample space and sigma algebra, which draws attention to the formal theory of probability (one where the above axioms hold) and applied probability, where the function  $P$  is supposed to model the probability of events in some real-world experiment.

**Example.** Consider two probability spaces,  $(\Omega, \mathcal{F}, P_1)$  and  $(\Omega, \mathcal{F}, P_2)$ , where  $\Omega = \{1, 2, 3, 4, 5, 6\}$  and  $\mathcal{F}$  is taken to be all subsets of  $\Omega$ . As previously noted, the function  $P$  is completely determined by the value of  $P(\omega_i)$  for each  $\omega_i \in \Omega$ . So define  $P_1(\omega) = \frac{1}{6}$  for all  $\omega \in \Omega$ , then  $P_1(\Omega) = \sum_{i=1}^6 P(\omega_i) = 1$ . This probability space is often associated with the probability of rolling a particular number on a fair dice, and is often referred to as an equiprobable function, or a uniform probability distribution - which are simply functions which assign the same probability measure to each outcome in the sample space. In contrast, define

$$P_2(\omega) = \begin{cases} \frac{1}{2} & \text{if } \omega = 1 \\ \frac{1}{8} & \text{if } \omega = 2, 3, 4, 5 \\ 0 & \text{if } \omega = 6 \end{cases}$$

Then again,  $P_2(\Omega) = \sum_{i=1}^6 P(\omega_i) = 1$ , but it is unclear what physical experiment this particular probability space may represent, even though it is a genuine probability space, since it is consistent with all of the axioms. This example also illustrates that  $P(\omega_i)$  can have the value zero without  $\omega_i$  being the empty set.

Also of key importance to our question is the concept of **independence of events**.

**Definition.** (Independent events) Let  $(\Omega, \mathcal{F}, P)$  be a probability space, and let  $A, B \in \mathcal{F}$  be events. Then  $A$  and  $B$  are said to be **independent** iff  $P(A \cap B) = P(A)P(B)$

Independence is a mathematical reflection of the notion that the occurrence of one event  $A$  has no effect on the occurrence of some other event  $B$ , and the other way around. In these cases, the resulting probability is the product of the individual probabilities for  $A$  and  $B$ . But the converse is true also, if, for example, it is discovered that the probability of two events taken together is equal to the product of their individual probabilities, we can conclude, at least in the Mathematical sense, that the two events are independent. In our particular question, we will assume the probability of flipping heads is  $1/2$  and that each coin flip is independent of the ones which proceed it.

We now turn to the question concerning the probability of flipping an infinite number of heads. Before we do so, we need to extend our definition of Probability to include sample spaces which are countably infinite. We will later extend this definition to also include sample spaces which are uncountably infinite.

**Definition** (Probability for countably infinite sample spaces). This definition is the same as our definition for the probability of finite sample spaces, with the addition that our sigma algebra may be countably infinite. Specifically, for every countably infinite collection of elements of  $A_k \in \mathcal{F}$ , we add the following

$$iv) P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k) \text{ if } A_i \cap A_j = \emptyset \text{ whenever } i \neq j$$

With this addition, it is possible to answer our question about the probability of an infinite toss of heads. However, it is perhaps more instructive to present an erroneous proof that the probability of an infinite toss of heads is 0.

*Erroneous Proof.* Assume that each toss of the coin is independent and the probability of getting heads on a single toss is  $\frac{1}{2}$ , then for  $k$  tosses, there are  $2^k$  distinct sets  $\{a_1, \dots, a_k\}$  such that  $a \in \{H, T\}$ , precisely one of which has  $a_i = H_i$  for  $1 \leq i \leq k$ . We denote this special case as  $A_k$  and define our countably infinite sample space to be,  $\Omega = \{A_1, A_2, \dots\}$  where  $\mathcal{F}$  is all subsets of  $\Omega$ , as usual. Define the probability  $P(A_k) = \frac{1}{2^k}$ , then  $P(A_k)$  represents the probability of getting  $k$  consecutive tosses of heads. Moreover, we have

$$P(\Omega) = P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k) = \sum_{k=1}^{\infty} \frac{1}{2^k} = 1$$

Now assume the probability of flipping an infinite number of heads is given by some positive number  $\epsilon$ , but then, by the Archimedean property  $\exists k \in \mathbb{N}$  such that  $0 < P(A_k) = \frac{1}{2^k} < \frac{1}{k} < \epsilon$ , but since  $P(A_k)$  represents the probability of flipping  $k$  heads in a row, it can't be less than the probability of flipping an infinite number of heads, hence the probability of flipping an infinite number of heads must be 0.  $\square$

As noted, there is a problem with this approach. It can be discovered by careful observation of our sample space  $\Omega$ . Observe,  $\Omega = \{A_1, A_2, A_3, \dots\} = \{\{H_1\}, \{H_1, H_2\}, \{H_1, H_2, H_3\}, \dots\}$ , but  $A_1 \cap A_2 = A_1 \neq \emptyset$ , and in general,  $A_k \cap A_{k+1} = A_k \neq \emptyset$ , hence  $P\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim_{k \rightarrow \infty} \frac{1}{2^k} = 0 \neq 1$  as earlier claimed.

To make this approach work for countably infinite samples spaces, mathematicians have introduced what are known as Bernoulli trials, where the  $A_k$  are not defined  $\{H_1, H_2, H_3, \dots, H_k\}$  but rather are made up of trials which continue until we flip a tail at which case the trial ends, so in our case,  $A_k$  would be interpreted the  $k^{th}$  flip that *first* results in a tail, so our Bernoulli trials look like  $A_1 = T_1$ ,  $A_2 = H_1T_2$ ,  $A_3 = H_1H_2T_3$ , and so on. Notice in this case each element of our sample space is  $A_k$  with the special property that each  $A_k$  represents the specified outcome of the  $k^{th}$  flip as being tails, so  $A_j \cap A_i = \emptyset$  for  $i \neq j$ . The limiting value of  $P(A_n)$  as  $n \rightarrow \infty$  is defined as the probability of flipping an infinite sequence of heads, which in this case is the same as  $\lim_{k \rightarrow \infty} \frac{1}{2^k} = 0$ . But we will leave this approach at the level of this outline, since our aim is to solve the problem where the event of flipping an infinite number of heads *is* included in the sample space, we will proceed along that line and leave the reader to explore that approach (see especially, [1, 3, 4]).

### III. “Randomly” choosing elements for infinite samples spaces.

To further motivate our discussion suppose we ask the question, What is the probability of randomly choosing a rational number in the set  $\Omega = (0, 1)$ ? Clearly our answer requires some sense of what it means to “randomly chose”

but more important for our discussion is the fact that  $\Omega$  is uncountably infinite, and our two previous definitions of a Probability measure do not include this case, so we will have to modify our approach one more time. Before doing so, we will prove a theorem that shows that for any uncountably infinite sample space  $\Omega$ , all but a countable number of events must have probability zero.

**Theorem 1.** *Let  $(\Omega, \mathcal{F}, P)$  be a probability space where  $\Omega$  is uncountably infinite, and let  $A$  be the set of all  $x \in \Omega$  with positive probability, then  $A$  is countable.*

*Proof.* For each  $n \in \mathbb{N}$ , define  $A_n = \left\{x \in \Omega \mid \frac{1}{n+1} < P(x) \leq \frac{1}{n}\right\}$ , then the size of  $A_n$  is at most  $n$ , since if its size were  $n+1$  or greater then the sum of all of the probabilities in  $A_n$  would be greater than  $(n+1) \left(\frac{1}{n+1}\right) = 1$ , which is impossible. Since every  $x \in A$  has positive probability, then  $x \in A_n$  for some  $n$ , and since  $A$  is the set of  $x \in \Omega$  with positive probability,  $A = A_1 \cup A_2 \cup \dots$ . Since the countable union of finite sets is countable,  $A$  is countable.  $\square$

What Theorem 1 tells us is that “some” elements of our set  $(0, 1)$  must have zero probability, since the cardinality of  $(0, 1)$  is uncountable, and only a countable subset of  $(0, 1)$  can have non-zero probability. But even if we chose  $S = \{x \in (0, 1) \mid x \in \mathbb{Q}\}$  to be the countable subset of  $(0, 1)$ , we can not define “randomly picked” to mean that each element of  $S$  is as equally probable as any other of being chosen.

**Theorem 2.** *Given a countably infinite sample space  $\Omega$ , then it is impossible to assign the same non-zero probability to every element in  $\Omega$ .*

*Proof.* Suppose to the contrary that every element of  $\Omega$  has the same probability of being picked, let  $\epsilon$  be that probability, so we have  $\sum_{i=1}^{\infty} \epsilon = 1$ , but since  $\epsilon$  is constant, by the Archimedean property,  $\exists n \in \mathbb{N}$  such that  $\frac{1}{n} < \epsilon$ , hence we have the contradiction,

$$1 = P(\Omega) = \sum_{i=1}^n \frac{1}{n} < \sum_{i=1}^n \epsilon < \sum_{i=1}^{\infty} \epsilon = 1$$

Hence it is impossible that every element of  $\Omega$  has the same probability of being chosen. <sup>2</sup>  $\square$

Observe that the above proof holds for all infinite sample spaces, since by Theorem 1, all but a countable number of outcomes can be given positive probability. Theorem 2 also tells us that when asking questions about picking elements at random from sample spaces which are infinite, some definition other than “each element has the same probability” must be used.

We mention the question of “picking at random” to introduce some of the difficulties that arise when talking about samples spaces which are infinite. We will not need this concept to answer our question about the probability of flipping a infinite sequence of heads, since in that case we will not be picking subsets of  $\Omega$  at random, but will rather pick subsets which have a particular property. What we will need, however, is to abandon assigning any positive probability to individual elements of  $\Omega$ . We develop this concept further below where, instead of individual elements, we consider intervals instead. First we need to give an informal definition of probability for uncountable sample spaces, where the sample space  $\Omega \subseteq \mathbb{R}$ .

<sup>2</sup>We could have equally shown that for any fixed positive number  $\epsilon$ , the infinite sum,  $\sum_{i=1}^{\infty} \epsilon$  is unbounded.

**Definition** (Probability for uncountable samples spaces). Let  $(\Omega, \mathcal{F}, P)$  be a probability space where  $\Omega \subseteq \mathbb{R}$  is uncountable and  $\mathcal{F}$  is a sigma-algebra of Borel subsets of  $\Omega$ , we will call  $A, B \in \mathcal{F}$ , measurable sets iff they the following properties:<sup>3</sup>

- i) If  $A$  is measurable, then so is  $\mathbb{R} \setminus A$ .
- ii) If  $\{A_1, A_2, \dots\}$  is a countable collection of measurable sets, then  $\{A_1 \cup A_2 \cup \dots\}$  is measurable.
- iii) Every interval (open, closed, half open, etc) is measurable.

Any non-negative, completely additive set function  $P$  defined on  $\mathcal{F}$  such that  $P(\Omega) = 1$  is a probability measure.

We can now set the stage for proving that the probability of flipping an infinite number of heads is zero. Our sample space  $\Omega$  will consist of sets of coin flips, such as  $\{H_1, T_2, T_3, H_4, \dots\}$  (the subscripts are used to distinguish consecutive coin flips), which are countably infinite in length. The start of one such set might look like  $\{H_1, T_2, T_3, H_4, T_5, T_6, H_7, \dots\}$  however, we will be interested in the subset of all such sets which represent at some point an infinite sequence of heads.

Now consider the well-know fact that in the interval  $(0, 1)$  the *decimal* expression of every irrational number consists of a never-repeating sequence of digits, and that every rational in the same interval comes in two forms, those with periods which repeat infinitely and those which terminate in all 0's. By also noting that every decimal expression that ends with all 0's can be replaced by an equivalent one which has all 9's (e.g.  $0.5 = 0.49999\dots$ ), our insight will be to realize that by looking at the base 2 representation of these numbers, we have infinite sequences of 0's and 1's that never repeat if the number is irrational, and for the rationals, we have sequences of 0's and 1's that forever repeat, and those which terminate with all 0's. By insisting that we represent these with all 1's instead (for example, in base 2,  $\frac{1}{2} = .1000\dots = .1111\dots$ ), and by observing that the set  $x \in (0, 1)$  expressed in base 2 by changing T to 1 and H to 0, is identical to our above set of all possible infinite sequences of coin flips, we have a model for the probability of flipping an infinite number of heads connected to the question of picking a subset of rational numbers from the interval  $(0, 1)$ . To exploit this connection, we will have to further develop the concept of measure as applied to subsets of the real line.

**Theorem 3.** *Let  $\mathbb{H}$  denote the set which contains all sets of consecutive coin flips which eventually result in an infinite sequence of Heads, then  $\mathbb{H}$  is countably infinite*

*Proof.* For the sake of simplicity, we will refer to the elements of  $\mathbb{H}$  as  $\{a_1, a_2, \dots\}$  where  $a \in \{H, T\}$ . Observe that every element  $x \in \mathbb{H}$  is a set which contains the result of the  $n^{\text{th}}$  coin toss and, since  $\mathbb{H}$  is the set of all possible coin tosses which contain an infinite sequence of heads, for some  $k$  all subsequent  $a_i$  for  $i \geq k$  have  $a = H$ .

To show that  $\mathbb{H}$  is infinite, just observe that for each  $n \in \mathbb{N}$ , there is an element of  $\mathbb{H}$  which has all tails before the  $n^{\text{th}}$  coin toss, and all heads thereafter.

To show that  $\mathbb{H}$  is countable, let 1 stand for heads, and 0 stand for tails, we can form a bijection between each element  $x \in \mathbb{H}$  and  $\mathbb{H}^*$  where each element in  $\mathbb{H}^*$  is the unique *binary* representation of some number in the interval

---

<sup>3</sup>For the sake of simplicity, we will not define what counts as a Borel set, but observe most subsets of  $\mathbb{R}$  one would normally consider is a Borel set. For a more extensive explanation, see [1, 7].

$(0, 1)$  which ends in a sequence of non-terminating 1's, and so represents an element of  $\mathbb{Q}$ . Since  $\mathbb{Q}$  is countable, so is  $\mathbb{H}$ .<sup>4</sup>  $\square$

With this in mind, we will be able to show that the probability of picking any rational from the interval  $(0, 1) = 0$ , which is actually a stronger statement than our original question, since the set  $\mathbb{H}$  is a bijection to a subset of  $\mathbb{Q}$ . In order to do so, we will need a few results from measure theory first.

**IV. The Measure of  $S = \{x \in (0, 1) \mid x \in \mathbb{Q}\}$ .** Since our probability will be defined on intervals, rather than individual elements, we need some sense of how this might be done. First observe that for our interval  $(0, 1)$ , we define  $P(a < X < b) = b - a$ . This notation may be confusing, but recall our outcomes are no longer individual elements of  $(0, 1)$  but rather subsets of  $(0, 1)$ , so  $P(a < X < b) = b - a$ , just says the probability of picking the entire interval  $(a, b)$  is just  $b - a$ . Our first step will be establish which subsets of  $(0, 1)$  have probability zero, or, using the vocabulary of Measure theory, which subsets of  $(0, 1)$  have measure zero.

**Definition.** A subset  $A \in \mathbb{R}$  is said to be a **zero set**, or have **measure zero**, if for every  $\epsilon > 0$ , there exists a countable collection of intervals  $(a_i, b_i)$  such that  $A \subseteq \bigcup_{i=1}^{\infty} (a_i, b_i)$  and  $\sum_{i=1}^{\infty} b_i - a_i \leq \epsilon$

We can use this definition to show that any finite subset of  $\mathbb{R}$  has measure zero.

**Theorem 4.** *Any given finite subset of  $\mathbb{R}$  has measure zero.*

*Proof.* Let  $A$  be a finite subset of  $\mathbb{R}$  that contains  $n$  elements, and denote the elements of  $A$  by  $\{x_1, x_2, \dots, x_n\}$ . Let  $\epsilon > 0$ , then  $A \subseteq \bigcup_{i=1}^n (x_i - \frac{\epsilon}{2n}, x_i + \frac{\epsilon}{2n})$  and  $\sum_{i=1}^n (x_i - \frac{\epsilon}{2n}, x_i + \frac{\epsilon}{2n}) = \sum_{i=1}^n \frac{2\epsilon}{2n} = \sum_{i=1}^n \frac{\epsilon}{n} \leq \epsilon$ , hence  $A$  has measure zero.  $\square$

What this means is that in some sense we can “corral off” finite subsets of  $\mathbb{R}$  and show that the probability of picking any finite subset of  $(0, 1)$  has Probability 0. This is an important consequence of our definition of a zero set, but since  $\mathbb{H}$  is countably infinite, we need to extend the results of Theorem 4 to countably infinite sets. Which is precisely what we do next.

**Theorem 5.** *Any countably infinite subset of  $\mathbb{R}$  has measure zero.*

*Proof.* Let  $A$  be a countably infinite subset of  $\mathbb{R}$  and denote the elements of  $A$  by  $\{x_1, x_2, \dots\}$ . Let  $\epsilon > 0$ , then  $A \subseteq \bigcup_{i=1}^{\infty} (x_i - \frac{\epsilon}{2^i}, x_i + \frac{\epsilon}{2^i})$ . Observe that the length of any given interval is  $\frac{\epsilon}{2^i} + \frac{\epsilon}{2^i} = \frac{2\epsilon}{2^i} = \frac{\epsilon}{2^{i-1}}$ , hence the length of all of the intervals is  $\sum_{i=1}^{\infty} \frac{\epsilon}{2^{i-1}} \leq \epsilon$ , hence  $A$  has measure zero.  $\square$

An easy corollary of Theorem 5 is the following:

**Corollary 6.** *The set  $\mathbb{Q}$  has measure zero.*

*Proof.* Since  $\mathbb{Q}$  is a countably infinite subset of  $\mathbb{R}$ , by Theorem 5,  $\mathbb{Q}$  has measure zero.  $\square$

---

<sup>4</sup>One obvious bijection is to observe for each  $h \in \mathbb{H}$  we have  $h = \{a_1, a_2, \dots\}$  where  $a \in \{H, T\}$ . Define  $f : \mathbb{H} \rightarrow \mathbb{R}$  by  $f(h) = \sum_{i=1}^{\infty} \frac{1}{2^i} (a_i)$  where  $a_i = 1$  if  $a = H$ , and 0 otherwise. To show that  $f$  is one to one and onto we leave to the reader.

While it is not necessary for the proof of our main theorem, we can now prove the following well-known result.

**Corollary 7.** *The set  $(0, 1) \setminus \mathbb{Q}$  has measure 1*

*Proof.* Observe that the measure of  $(0, 1)$  is  $1 - 0 = 1$ . Since the measure of two disjoint sets is the sum of their individual measures<sup>5</sup>, then by Corollary 6, the set  $(0, 1) \setminus \mathbb{Q}$  has measure 1.  $\square$

We can now prove that the probability of flipping an infinite sequence of heads is zero.

**Theorem 8.** *The probability of flipping an infinite sequence of heads is 0.*

*Proof.* Let  $(\Omega, \mathcal{F}, P)$  be a probability space where  $\Omega = (0, 1)$  and suppose  $\mathcal{F}$  is a sigma-algebra of Borel subsets of  $\Omega$ . Let  $\mathbb{H}^*$  be the subset of  $\Omega$  as defined in Theorem 3, then  $P(\mathbb{H}^*) = 0$ , since  $\mathbb{H}^* \subset \mathbb{Q}$ , by relating  $\mathbb{H}^*$  to  $\mathbb{H}$ , we have that the probability of flipping an infinite sequence of heads is zero.  $\square$

#### REFERENCES

- [1] Tom M Apostol. *Calculus*, volume II. Blaisdell Pub, 1961.
- [2] Santiago Cañez. The riemann-lebesgue theorem. <http://math.berkeley.edu/~scanez/courses/math104/fall11/handouts/riemann-lebesgue.pdf>.
- [3] William Feller. *An Introduction to Probability Theory and Its Applications*. Wiley, second edition, 1957.
- [4] Richard Isaac. *The Pleasures of Probability*. Springer-Verlag, 1995.
- [5] Frank Jones. *Lebesgue Integration on Euclidean Space*. Jones and Bartlett, 1993.
- [6] Terence Tao. *An Introduction to Measure Theory*. American Mathematical Society, 2011.
- [7] Warren Weaver. *Lady Luck : The Theory of Probability*. Dover, 1982.

SAN ANTONIO COLLEGE

*E-mail address:* `kslinker@alamo.edu`

---

<sup>5</sup>We have not proven this, but the reader may consult [5, 2, 6] for elementary proofs.