Probability and Measure Solutions

WispyAbyss

December 10, 2024

Contents

Forward

This document will contain notes and solutions corresponding to Probability and Measure, Third Edition, by Patrick Billingsley [amazon].

Chapter 1.1 - Borel's Normal Number Theorem

Notes

For a complete understanding of probability, you need to understand an infinite number of events as well as a finite number of events. We try and present why that must be so here.

The Unit Interval

We take the length of an interval $I = (a, b] = b - a$. Note, for A a disjoint set of intervals in $(0, 1]$, we have that $P(A)$ is well defined. If B is a similar disjoint set, and is disjoint from A, $P(A+B) = P(A) + P(B)$ is well defined as well. Note - we haven't defined anything for intersections yet. These definitions can also directly stem from the Riemann integral of step functions.

The unit interval can give the probability that a single particle is emitted in a unit interval of time. Or a single phone call comes in. However, it can also model an infinite coin toss. This is done as follows - for $\omega \in (0, 1]$, define:

$$\omega = \sum_{n=1}^{\infty} \frac{d_n(\omega)}{2^n}$$

Where $d_n(\omega)$ is 0 or 1, and comes from the binary expansion of ω. We take ω as the non terminating representation. Note, we were particular when we defined intervals as half inclusive. Examine the set of ω for which $d_i(\omega) = u_i$ for $i = 1, \cdots, n$, $u_i \in \{0, 1\}$. We have that:

$$\sum_{i=1}^{n} \frac{u_i}{2^i} < \omega \leq \sum_{i=1}^{n} \frac{u_i}{2^i} + \sum_{i=n+1}^{\infty} \frac{1}{2^i}$$

We cannot have the lower extreme value, as this would imply ω takes on its terminating binomial representation, which is what we said we would not do. This is our first taste, I guess, of measure 0 sets, we we still have:

$$\mathbb{P}\left[\omega : d_i(\omega) = u_i, i = 1, \cdots, n\right] = \frac{1}{2^n}$$

Note, probabilities of various familiar events can be written down immediately. Ultimately, note, however, each probability is the sum of disjoint dyadic intervals of various ranks k. Ie, all the events are still well defined by our probability definition above. We have:

$$\mathbb{P}\left[\omega : \sum_{i=1}^{n} d_i(\omega) = k\right] = \binom{n}{k} \frac{1}{2^n}$$

All these results have been for finitely many components of $d_i(\omega)$. What we are interested in, however, is properties of the entire sequence of $\omega = (d_1(\omega), d_2(\omega), \cdots)$.

The Weak Law of Large Numbers

What I like about this chapter, is to me - it *emphasizes* the connection between the *structure of real numbers*, and probability. At the end of the day - probability can be seen as just extracting properties of *frequency* over the real numbers, to be understood as probabilistic statements. However, with just our basic real numbers - we can't really prove a lot of properties about infinite things. That is when measure theory comes in later. However, for now, we look at what we can prove - and that starts with the weak law of large numbers. We have:

Theorem 1.1 - The Weak Law of Large Numbers For each ϵ:

$$\lim_{n\to\infty} \mathbb{P}\left[\omega : \left|\frac{1}{n}\sum_{i=1}^{n} d_i(\omega) - \frac{1}{2}\right| \geq \epsilon\right] = 0$$

Probabilistically - this is saying that if n is large, then there is a small probability that the fraction/relative *frequency* of heads in n tosses will deviate much from 1/2. Think about it as a statement over the real numbers as well - it is also interesting. Ultimately, the intervals containing ω that do not satisfy the above are getting smaller and smaller and smaller. We formalize this with the following concept:

As $d_i(\omega)$ are constant over each dyadic interval of rank n if $i \leq n$, the sums $\sum_{i=1}^{n} d_i(\omega)$ are also constant over rank n. Thus, the set in the theorem is just a disjoint union of dyadic intervals of rank n. Note - the theorem is saying, that the total weight given to those intervals gets smaller and smaller as n goes to infinity.

Now, we go over how to prove the theorem. It relies on rademacher variables:

$$r_n(\omega) = 2d_n(\omega) - 1$$

These are ± 1 when $d_n = 1/0$. Note, these have the same "being constant on dyadic intervals" properties as $d_n(\omega)$. We define:

$$s_n(\omega) = \sum_{i=1}^{n} r_i(\omega)$$

And so, our theorem is equivalent to proving:

$$\lim_{n\to\infty} \mathbb{P}\left[\omega : \left|\frac{1}{n}s_n(\omega)\right| \geq \epsilon\right] = 0$$

Note, rademacher functions also have interpretations, probabilistically, of random walks and such. With these variables, we can ultimately find properties, going all the way to:

$$\int_0^1 s_n^2(\omega) = n$$

However, what interests me is the following: Chebyshev's Lemma, but as a property of the real numbers. We have:

Lemma - Chebyshev's Inequality If f is a nonnegative step function, then $[\omega : f(\omega) \geq \alpha]$ is for $\alpha > 0$ a finite union of intervals, and:

$$\mathbb{P}\left[\omega : f(\omega) \geq \alpha\right] \leq \frac{1}{\alpha} \int_0^1 f(\omega) d\omega$$

Proof: Note, it is all just properties of step functions. Let c_j correspond to the step intervals $(x_{j-1}, x_j]$, and let $\sum'$ be the sum over $c_j \geq \alpha$. Then, we have quite easily:

$$\int_0^1 f(\omega) d\omega = \sum c_j (x_j - x_{j-1}) \geq \sum{}' c_j (x_j - x_{j-1}) \geq \sum{}' \alpha (x_j - x_{j-1}) = \alpha \mathbb{P}\left[\omega : f(\omega) \geq \alpha\right]$$

Thus, we have Chebyshev's inequality, and with it, we can easily prove the Weak Law of Large Numbers. However - it is important to note - these are *properties over the real numbers*, as much as they are probabilistic properties.

Solutions