

Probability Theory and Measure: Lecture V

Charles B. Moss

August 31, 2026

Old White Pass



Old White Pass Bridge



Reading Question

- A Frequency Approach
- B Personal Probability
- C Axiomatic Probabilities
 - 1 Andrei Kolmogorov
 - 2 Leonard Savage and Bruno de Finetti
 - 3 Huygens

- 1 Uniform Probability Measure
 - Lebesgue Measure
- 2 Lebesgue Measure and Lebesgue Integral
- 3 Random Variables and Distributions

Uniform Probability Measure

- I think that Bieren's discussion of the uniform probability measure provides a firm basis for the concept of probability measure.
 - First, we follow the conceptual discussion of placing ten balls numbered 0 through 9 into a container. Next, we draw out an infinite sequence of balls out of the container, replacing the ball each time.
 - In Excel, we can mimic this sequence using the function `floor(rand()*10,1)`. This process will give a sequence of random numbers such as presented in Table 1.

Table: Random Draws of Single Digits

Ball Drawn	Draw 1	Draw 2	Draw 3
1	7	0	3
2	4	2	0
3	1	9	2
4	4	6	2
5	8	4	0
6	3	5	4

- Taking each column, we can generate three random numbers 0.741483, 0.029645, 0.302204. Note that each of these sequences are contained in the unit interval $\Omega = [0, 1]$. The primary point of the demonstration is that the number draw $\{x \in \Omega = [0, 1]\}$ is a probability measure.
 - Taking $x = 0.741483$ as the example, we want to prove that $P([0, x = 0.741483]) = 0.741483$. To do this we want to work out the probability of drawing a number less than 0.741483.
 - As a starting point, what is the probability of drawing the first number in Table 1 less than 7, it is $7 \sim \{0, 1, 2, 3, 4, 5, 6\}$. Thus, without consider the second number, the probability of drawing a number less than 0.741483 is somewhat greater than $7/10$.

- Next, we consider drawing a second number given that the first number drawn is greater than or equal to 7. Now, we are interested in the scenario where the number drawn is equal to seven. This occurs 1/10 of the times. Note that the two scenarios are disjoint. If the first draw is less than seven, it is not equal to seven. Thus, we can rely on the summation rule of probabilities

$$\text{If } A_i \cap A_j = \emptyset \text{ then } P \left[\bigcup_{k=1}^n A_k \right] = \sum_{k=1}^n P[A_k] \quad (1)$$

- Thus, the probability of drawing a number less than 0.74 is the sum of drawing the first number less than 7 and the second number less than 4 given that the first number drawn is 7. The probability of drawing the second number less than 4 is $4/10 \sim \{0, 1, 2, 3\}$. Given that the first number equal to 7 only occurs $1/10$ of the time, the probability of the two events is

$$P([0, x = 0.74]) = \frac{7}{10} + \frac{4}{10} \left(\frac{1}{10} \right) = \frac{7}{10} + \frac{4}{100} = 0.74 \quad (2)$$

- Continuing to iterate this process backward, we find that $P([0, x = 0.741483]) = 0.741483$. Thus, for $x \in [0, 1]$ we have $P([0, x]) = x$.

- A couple of things, first following our last lecture we have justified the use of probability measure defined on a σ -algebra. More concretely, we can define the unit interval as a special form of a σ -algebra called a Borel set. Second, the probability structure generated is a commonly used probability function, the uniform random variable, $U[0, 1]$.
 - Definition 1.5: The σ -algebra generated by the collection

$$C = \{ (a, b) : \forall a < b, a, b \in \mathbb{R} \} \quad (3)$$

of all open intervals in $\mathbb{R}$ is called the Euclidean Borel field, denoted B , and its members are called Borel sets.

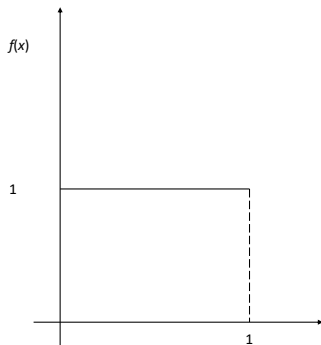
- In this case, we have defined $a = 0$ and $b = 1$.

- Also, note that for any $x \in [0, 1]$, $P(\{x\}) = P([x, x]) = 0$. This has the advantage of eliminating the lower end of the range. Specifically,

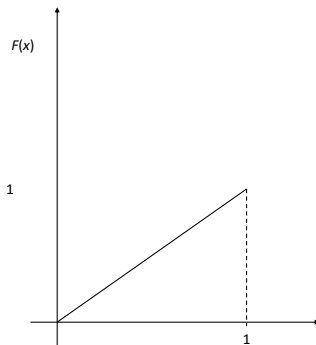
$$P([0, x]) = P([0]) + P((0, x)) \quad (4)$$

Further, we have that as long as $a < b$, $a, b \in [0, 1]$

$$P([a, b]) = P((a, b]) = P([a, b)) = P((a, b)) = b - a \quad (5)$$



Probability Density Function



Cumulative Density Function

Reimann Sums

$$\begin{aligned} S_1(K) &= \sum_k |x_k - x_{k-1}| \sup_{[x_{k-1}, x_k]} f(x) \\ S_2(K) &= \sum_k [x_k - x_{k-1}] \inf_{[x_{k-1}, x_k]} f(x) \end{aligned} \tag{6}$$

- In the limit

$$\lim_{K \rightarrow \infty} \epsilon(K) = S_1(K) - S_2(K) = 0 \tag{7}$$

Lebesgue Measure

- In the Bieren's formulation

$$F_0 = \{ (a, b) : [a, b], (a, b], [a, b), \forall a, b \in [0, 1], a < b, \text{ and their countable union} \} \quad (8)$$

This is probability measure is a special case of the Lebesgue measure.

Lebesgue Measure and Lebesgue Integral

- Building on the Uniform Distribution example above, we next define the Lebesgue measure as a function λ that measures the length of the interval (a, b) on any Borel set B in $\mathfrak{R}$

$$\lambda(B) = \inf_{B \subset \bigcup_{j=1}^{\infty} (a_j, b_j)} \sum_{j=1}^{\infty} \lambda((a_j, b_j)) = \inf_{B \subset \bigcup_{j=1}^{\infty} (a_j, b_j)} \sum_{j=1}^{\infty} (b_j - a_j) \quad (9)$$

It is the total length of the Borel set taken from the outside.

- Based on the Lebesgue measure, we can then define the Lebesgue integral based on the basic definition of the Riemann integral:

$$\int_a^b f(x) dx = \sup \sum_{m=1}^n \left(\inf_{x \in I_m} f(x) \right) \lambda(I_m) \quad (10)$$

- Replacing the interval of the summation, the Lebesgue integral becomes:

$$\int_A f(x) dx = \sup \sum_{m=1}^n \left(\inf_{x \in B_m} f(x) \right) \lambda(B_m) \quad (11)$$

Random Variables and Distributions

- Now we have established the existence of a probability measure based on a specific form of σ -algebras called Borel fields. The question is then can we extend this rather specialized formulation to broader groups of random variables? Of course, or this would be a short class.
- As a first step, let's take the simple coin-toss example. In the case of a coin there are two possible outcomes (heads or tails). These outcomes completely specify the sample space.
 - To add a little structure, we construct a random variable X that can take on two values $X = 0$ or 1 .
 - If $X = 1$ the coin toss resulted in a heads while if $X = 0$ the coin toss resulted in a tails.
 - Next, we define each outcome based on an event space ω :

$$\begin{aligned} P(X = 1) &= P(\{\omega \in \Omega : X(\omega) = 1\}) = P([H]) \\ P(X = 0) &= P(\{\omega \in \Omega : X(\omega) = 0\}) = P([T]) \end{aligned} \quad (12)$$

- The probability function is then defined by the random event ω . Defining ω as a uniform random variable from our original example one alternative would be to define the function as

$$X(\omega) = 1 \text{ if } \omega \leq 0.50 \quad (13)$$

This definition results in the standard 50-50 result for a coin toss. However, it admits more general formulations. For example, if we let

$$X(\omega) = 1 \text{ if } \omega \leq 0.40 \quad (14)$$

The probability of heads becomes 40 percent.

Readings

- Moss 3.1 - 3.4
- Popper – Chapter 2