Math Colloquium
Five Lectures on Probability Theory
Part 4: Simple Random Walk

Robert Niedzialomski, rniedzia@utm.edu, utm.edu/staff/rniedzia

March 17th, 2021

Kolmogorov Axioms of Probability

Let Ω be a set. A collection $\mathfrak{F}$ of subsets of Ω is called of σ-**algebra** of if

- $\emptyset, \Omega \in \mathfrak{F}$,
- $A, B \in \mathfrak{F} \Longrightarrow A - B \in \mathfrak{F}$.
- $A_n \in \mathfrak{F}, n \in \mathbb{N} \Longrightarrow \cup_{n=1}^{\infty} A_n \in \mathfrak{F}$, and $\cap_{n=1}^{\infty} A_n \in \mathfrak{F}$

A **probability measure** on a measurable space $(\Omega, \mathfrak{F})$ is a mapping $P \colon \mathfrak{F} \to [0, 1]$ satisfying

- $P(\emptyset) = 0$ and $P(\Omega) = 1$.
- $A_n \in \mathfrak{F}$, $n \in \mathbb{N}$, disjoint sets $\Longrightarrow P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$.

The triple $(\Omega, \mathfrak{F}, P)$ is called a **probability space**.

Examples of σ-algebras

($\bullet$) The collection 2^Ω of all subsets of Ω is a σ-algebra.

($\bullet$) If $A \subset \Omega$ with $A \neq \emptyset$ and $A \neq \Omega$, then $\{\emptyset, A, \Omega - A, \Omega\}$ is a σ-algebra.

($\bullet$) For any collection $\mathfrak{A} \subset 2^\Omega$ of subsets of Ω there exists the smallest σ-algebra containing $\mathfrak{A}$, denoted by $\sigma(\mathfrak{A})$. The construction of $\sigma(\mathfrak{A})$: Define

$$\mathfrak{B} = \{\mathfrak{F} \subset 2^\Omega \colon \mathfrak{A} \subset \mathfrak{F} \quad \text{and} \quad \mathfrak{F} \text{ is a } \sigma\text{-algebra}\}.$$

and put

$$\sigma(\mathfrak{A}) = \bigcap_{\mathfrak{F} \in \mathfrak{B}} \mathfrak{F}$$

($\bullet$) Let Ω be a topological space. The smallest σ-algebra $\mathcal{B}(\Omega)$ containing all open (hence closed) sets is called the **Borel** σ-algebra.

Random Variables

Let $(\Omega, \mathfrak{F}, P)$ be a probability space. A function $X\colon \Omega \to \mathbb{R}^d$ is called a **random variable** is for any Borel set $B \in \mathcal{B}(\mathbb{R}^d)$ we have

$$X^{-1}(B) = \{\omega \in \Omega\colon X(\omega) \in B\} \in \mathfrak{F}.$$

The collection $\{X^{-1}(B)\colon B \in \mathcal{B}\}$ is a σ-algebra of subsets of Ω denoted by $\sigma(X)$.

A random variable X defines a probability measure $P_X = P \circ X^{-1} = X\sharp P$ on the measurable space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ as follows

$$P_X(B) = P(X^{-1}(B)) = P(X \in B).$$

We call P_X the **probability distribution** of X (or the **law** of X or the **push-forward** of P via X)

Integration

Let $(\Omega, \mathfrak{F}, P)$ be a probability space and let $X\colon \Omega \to [0,\infty]$ be is a random variable of the form

$$X = \sum_{j=1}^{m} c_j I_{A_j},$$

where $c_j \geq 0$, $A_j \in \mathfrak{F}$, and I_A denotes the characteristic function of a set $A \subset \Omega$ and is defined by: $I_A(\omega) = 1$ when $\omega \in A$ and $I_A(\omega) = 0$ when $\omega \notin A$. Such random variables are called **simple**. The **integral** of X is defined by

$$\int_\Omega X(\omega)\, dP(\omega) = \sum_{j=1}^{m} c_j P(A_j).$$

For any non-negative random variable $Y\colon \Omega \to [0,\infty]$ we let

$$\int_\Omega Y(\omega)\, dP(\omega) = \sup\{\int_\Omega X\, dP\colon X \leq Y \quad \text{and} \quad X \text{ is simple}\}$$

For any random variable $X\colon \Omega \to [-\infty, \infty]$ we first decompose

$$X^+ = \max(0, X),\; X^- = -\min(0, X) \Rightarrow X = X^+ - X^- \text{ and } X^+, X^- \geq 0$$

and then define

$$\int_\Omega X(\omega)\, dP(\omega) = \int_\Omega X^+(\omega)\, dP(\omega) - \int_\Omega X^-(\omega)\, dP(\omega).$$

Some Properties of Integration:

- If X is a non-negative random variable with $P(X = \infty) > 0$, then $\int_\Omega X(\omega)\, dP(\omega) = \infty$
- For non-negative random variables $X_1, X_2, \ldots$ we have that $\int_\Omega \sum_{n=1}^\infty X_n(\omega)\, dP(\omega) = \sum_{n=1}^\infty \int_\Omega X_n(\omega)\, dP(\omega)$
- If X is a non-negative random variable such that $\int_\Omega X(\omega)\, dP(\omega) = 0$, then $X = 0$ almost surely, which means that $P(X = 0) = 1$.

Expectation and Variance

Let X be a random variable on a probability space $(\Omega, \mathfrak{F}, P)$. Its **expectation** (or **mean**) is defined by

$$E[X] = \int_{\mathbb{R}} x \, dP_X(x) = \int_{\Omega} X(\omega) \, dP(\omega).$$

Hence expectation of a random variable is just its integral. The **variance** of X is

$$\mathrm{Var}[X] = E[|X - E[X]|^2] = E[X^2] - E[X]^2$$

Example

Let $m \in \mathbb{R}$ and let $\sigma^2 > 0$. The **normal distribution** with mean m and variance σ^2 is the probability measure N_{m,σ^2} on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ given by

$$N_{m,\sigma^2}(B) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_B e^{-\frac{(x-m)^2}{2\sigma^2}} \, dx$$

Independence

Let $(\Omega, \mathfrak{F}, P)$ be a probability space. Two random variables $X\colon \Omega \to \mathbb{R}$ and $Y\colon \Omega \to \mathbb{R}$ are **independent** if for any two Borel sets $A \subset \mathbb{R}$ and $B \subset \mathbb{R}$ the events $(X \in A)$ and $(X \in B)$ are independent, i.e.,

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B).$$

Lemma

If two random variables X and Y independent, then $E[XY] = E[X]E[Y]$.

Lemma

If two random variables X and Y defined on some discrete probability space are independent, then $Var(X + Y) = Var(X) + Var(Y)$.

Random Walk

Let $(\Omega, \mathfrak{F}, P)$ be a probability space and let $(S, \mathcal{S})$ be a measurable space, called the **state space**. A sequence $(X_n)_{n=1}^{\infty}$ of random variables $X_n \colon \Omega \to S$ is called a **discrete time stochastic process**.

Let $(X_n)_{n=1}^{\infty}$ be a discrete time stochastic process with the state space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Suppose that the random variables X_n are independent and identically distributed. Then the sequence $(S_n)_{n=1}^{\infty}$ of partial sums, i.e,

$$S_n = \sum_{i=0}^{n} X_i$$

is called a **random walk** starting at $X_0 = 0$.
If $x \in \mathbb{R}^d$, then the sequence

$$S_n = x + \sum_{i=1}^{n} X_i$$

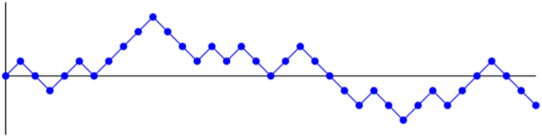
is called a **random walk** starting at $X_0 = x$.

Simple Random Walk

A **Simple Random Walk** is a sequence (S_n) of partial sums associated to the independent identically distributed random variables X_n with values in $\mathbb{R}^d$ and distribution

$$X_n = \pm e_i \quad \text{with probability } \frac{1}{2d}$$

Simple Random Walk in 1D: When $d = 1$, X_n is the Bernoulli distribution, i.e., X_n takes values -1 and 1 with probability $\frac{1}{2}$. We think of jumping one unit to the right when $X_n = 1$ and one unit to the left when $X = -1$. This way we have a random walker that starts his/her walk at the origin. We visualize that walk by the graph (n, S_n).



Two Fundamental Questions

Question 1: With what probability does the random walker ever return to the starting position?

Question 2: How often (on average) does the random walker return to the starting position?

To answer these two questions we define the following random variable. We put

$$V = \sum_{n=0}^{\infty} I_{\{S_{2n}=0\}}$$

Then V takes the values $k = 1, 2, \ldots, \infty$ and it measures how many times the random walker returns to the origin.

Random Walk Continued

Let q be the probability that the random walker ever returns to the origin. Then we will show that

$$P(V = k) = q^{k-1}(1 - q) \quad \text{for any } k = 1, 2, \ldots.$$

Hence,

$$\begin{aligned} E[V] &= \infty \cdot P(V = \infty) + \sum_{k=1}^{\infty} kP(V = k) \\ &= \infty \cdot P(V = \infty) + \sum_{k=1}^{\infty} kq^{k-1}(1 - q) \end{aligned}$$

where we use the convention $\infty \cdot 0 = 0$.
Moreover one can show that (we did it last time in dimension $d = 1$)

$$P(S_{2n} = 0) \sim \frac{c_d}{n^{d/2}} \quad \Rightarrow \quad E[V] = \begin{cases} \infty, & d = 1, 2 \\ < \infty, & d \geq 3 \end{cases}$$

Case 1: Assume $q = 1$.

Then $P(V = k) = 0$ for any $k = 1, 2, \ldots$. This implies that $P(V = \infty) = 1$. Thus the random walker returns to the origin infinitely many times with probability 1.

Case 2: Assume $q < 1$.

Then

$$P(V < \infty) = \sum_{k=1}^{\infty} P(V = k) = \sum_{k=1}^{\infty} q^{k-1}(1 - q) = 1$$

Hence the random walker returns to the origin only finitely many times with probability one (and $P(V = \infty) = 0$).

Polya's Theorem

Theorem (Polya)

In dimensions 1 *and* 2 *the random walker returns to the origin infinitely many times with probability one. We say that the random walk is* **recurrent**. *In dimensions* $d \geq 3$ *the random walker returns to the origin only finitely many times with probability one. We say that the random walk is* **transient**.

Proof: Let $d = 1, 2$. We will show that $q = 1$. Suppose, by contradiction, that $q < 1$. Then $P(V = \infty) = 0$. Therefore, $E[V] = \frac{1}{1-q} < \infty$. Contradiction.

Let $d \geq 3$. Then $E[V] < \infty$. Therefore $P(V = \infty) = 0$. We will show $q < 1$. Suppose, by contradiction, that $q = 1$. Then $P(V = \infty) = 1$. Contradiction.

Stopping Time

We have been analyzing the random variable V defined by

$$V = \sum_{n=0}^{\infty} I_{\{S_{2n}=0\}}.$$

We defined

- the walk is recurrent exactly when $P(V = \infty) = 1$
- the walk is transient exactly when $P(V < \infty) = 1$.

Let T denote the first time the walker returns to the origin, i.e., let T be the random variable

$$T = \min\{2n \colon S_{2n} = 0\}.$$

We have shown that (since $P(T < \infty) = q$)

- the walk is recurrent exactly when $P(T < \infty) = 1$
- the walk is transient exactly when $P(T < \infty) < 1$.

Lemma

For any $k = 1, 2, \ldots$ *we have*

$$P(V = k) = (1 - q)q^{k-1}$$

Proof: First we will show that

$$P(V = k + 1) = P(V = k)P(T < \infty), \quad k = 1, 2, \ldots \tag{1}$$

Suppose the first return to the origin occurs at time $T = j$. Then $V = k + 1$ implies the random walk $S'_n = S_{n+j} - S_j$ makes k returns to the origin. By independence

$$P(V = k + 1, T = j) = P\left(\sum_{n=0}^{\infty} I_{\{S'_{2n}=0\}} = k, T = j\right)$$

$$= P\left(\sum_{n=0}^{\infty} I_{\{S'_{2n}=0\}} = k\right) P(T = j) = P(V = k)P(T = j).$$

Summation over all j gives identity (1). Now proceed by induction.

References

- Lawler, Gregory F. Random walk and the heat equation. Student Mathematical Library, 55. American Mathematical Society, Providence, RI, 2010