Theoretical Statistics

STAT 210A

Instructor: Will Fithian

Kelvin Lee

UC Berkeley

Contents

Lecture 1

Measure Theory

1.1 Basics

1.1.1 Measures

Definition 1.1.1 (Measure)**.** Given a set $\mathcal{X}$, a **measure** μ maps subsets $A \subseteq \mathcal{X}$ to nonnegative numbers $\mu(A) \in [0, \infty)$.

Example 1.1.2. Let $\mathcal{X}$ be a countable set ($\mathcal{X} = \mathbb{Z}$ for example). Then the **counting measure** ia

$$\mu(A) = \#A = \# \text{ of points in } A.$$

Example 1.1.3. Consider $\mathcal{X} = \mathbb{R}^n$. The **Lebesgue measure** is

$$\lambda(A) = \int \cdots \int_A dx_1 \cdots dx_n = \text{Vol}(A).$$

Example 1.1.4 (Standard Gaussian Distribution)**.**

$$\mathbb{P}(A) = \mathbb{P}(Z \in A) = \int \cdots \int_A \phi(x) dx_1 \cdots dx_n$$

where $Z \sim \mathcal{N}_n(0, I_n)$ and $\phi(x) = \frac{e^{-\frac{1}{2}\sum x_i^2}}{\sqrt{(2\pi)^n}}$

Because of pathological sets, $\lambda(A)$ is only defined for some subsets $A \subseteq \mathbb{R}^n$. In other words, it is often impossible to assign measures to all subsets A of $\mathcal{X}$. This leads to the idea of a σ-field(σ-algebra).

Definition 1.1.5 (σ-field)**.** A **σ-field** is a collection of sets on which μ is defined, satisfying certain closure properties.

In general, the domain of a measure μ is a collection of subsets $\mathcal{F} \subseteq 2^{\mathcal{X}}$ (power set), and $\mathcal{F}$ must be a σ-field.

Example 1.1.6. Let $\mathcal{X}$ be a countable set. Then $\mathcal{F} = 2^{\mathcal{X}}$. (Counting measure is defined for all subsets).

Example 1.1.7. Let $\mathcal{X} = \mathbb{R}^n$, then $\mathcal{F}$ is the **Borel σ-field** $\mathcal{B}$, the smallest σ-field containing all open rectangles $(a_1, b_1) \times \cdots \times (a_n, b_n)$ where $a_i < b_i \quad \forall i$.

Given a **measurable space** $(\mathcal{X}, \mathcal{F})$, a **measure** is any map $\mu : \mathcal{F} : [0, \infty)$ with $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$ if $A_i \in \mathcal{F}$ are disjoint. If $\mu(\mathcal{X}) = 1$, then μ is a **probability measure**.

Measures let us define integrals that put weight $\mu(A)$ on $A \subseteq \mathcal{X}$.

Define

$$\int \mathbf{1}\{x \in A\} d\mu(x) = \mu(A) \qquad \text{(indicator)}$$

extend to other functions by linearity and limits:

$$\int \left(\sum c_i \mathbf{1}\{x \in A_i\}\right) d\mu(x) = \sum c_i \mu(A_i) \qquad \text{(simple function)}$$

$$\int f(x) d\mu(x) \qquad \text{((measurable) function approx. by simple functions)}$$

Example 1.1.8.

- *Counting*: $\int f d\# = \sum_{x \in \mathcal{X}} f(x)$.
- *Lebesgue*: $\int f d\lambda = \int \cdots \int f(x) dx_1 \cdots dx_n$.
- *Gaussian*: $\int f dP = \int \cdots \int f(x)\phi(x) dx_1 \cdots dx_n = \mathbb{E}[f(Z)]$.

1.1.2 Densities

The λ and $\mathbb{P}$ above are closely related and we now want to make this precise.

Given $(\mathcal{X}, \mathcal{F})$, two measures $\mathbb{P}, \mu$, we say that $\mathbb{P}$ is **absolutely continuous** with respect to μ if $\mathbb{P}(A) = 0$ whenever $\mu(A) = 0$.

Notation: $P \ll \mu$ or we say μ *dominates* $\mathbb{P}$.

If $\mathbb{P} \ll \mu$, then (under mild conditions) we can always define a **density function** $\rho : \mathcal{X} \to [0, \infty)$ with

$$\mathbb{P}(A) = \int_A \rho(x) d\mu(x)$$
$$\int f(x) d\mathbb{P}(x) = \int f(x)\rho(x) d\mu(x).$$

The density function is also defined as

$$\rho(x) = \frac{d\mathbb{P}}{d\mu}(x),$$

known as *Radon-Nikodyan derivative*.

Remark. It is useful to turn $\int f d\mathbb{P}$ into $\int f \rho d\mu$ if we know how to calculate integrals $d\mu$.

If $\mathbb{P}$ is a probability measure, μ is a Lebesgue measure, then $\rho(x)$ is called **probaility density function** (pdf). If μ is a counting measure, then $\rho(x)$ is called the **probability mass function** (pmf).

1.1.3 Probability Space and Random Variables

Typically, we set up a problem with multiple random variables having various relationships to one another. It is convenient to think of them as functions of an abstract outcome ω.

Definition 1.1.9. Let $(\omega, \mathcal{F}, \mathbb{P})$ be a **probability space**. $\omega \in \Omega$ is called **outcome**. $A \in \mathcal{F}$ is called **event**. $\mathbb{P}(A)$ is called **probability of** A.

Definition 1.1.10. A **random variable** is a function $X : \Omega \to \mathcal{X}$. We say $\mathcal{X}$ has distribution Q, denoted as $X \sim Q$ if $\mathbb{P}(X \in B) = \mathbb{P}(\{\omega : X(\omega) \in B\}) = Q(B)$.

More generally, we could write events involving many random variables.

Definition 1.1.11. The **expectation** is an integral with respect to $\mathbb{P}$:

$$\mathbb{E}[f(X,Y)] = \int_\Omega f(X(\omega), Y(\omega)) d\mathbb{P}(\omega) = .$$

To do real calculations, we must eventaully boil $\mathbb{P}$ or $\mathbb{E}$ down to concrete integrals/sums/etc. If $\mathbb{P}(A) = 1$, we say that A occurs **almost surely**.

Lecture 2