Statistics 210A – Theoretical Statistics

Albert Zhang
Professor Will Fithian

Fall 2019

Contents

1 August 29

1.1 Measure Theory Basics

Given a set X, a measure μ maps subses $A \subseteq X$ to $[0, \infty]$.

Example 1.1

Let X be a countable set, such as $\mathbb{Z}$. Define the counting measure as $\#(A) = |A|$.

Example 1.2 (Lebesgue Measure)

Define the Lebesgue measure $\lambda(A) =$ "volume of A".

Consider the (standard) normal distribution. $P(A) := P(Z \in A)$, where $Z \sim \mathcal{N}(0, 1)$.

Remark. Because of pathologies, $\lambda(A)$ can only be defined for certain subsets of $A \subseteq \mathbb{R}^n$. In general, domain of a measure is a σ-field $\mathcal{F} \in 2^X$ satisfying certain closure properties. If X is countable, we may just take $\mathcal{F} = 2^X$. If $X = \mathbb{R}^n$, we take $\mathcal{F} = \mathcal{B}^n$, the borel σ-field generated by all open rectangles $(a_1, b_1) \times \cdots \times (a_n, b_n)$, $a_i < b_i$.

Given a measurable space $(X, \mathcal{F})$, a measure is a map $\mu : \mathcal{F} \to [0, \infty]$ satisfying

$$\mu \left(\cup_{i=1}^n A_i \right) = \textstyle\sum_{i=1}^n \mu(A_i) \text{ for disjoint sets } A_1, A_2, \ldots, A_n \in \mathcal{F}.$$

Example 1.3 (Probability Measure)

A probability measure P is a measure with $P(X) = 1$.

Measures let us define integrals that put weight $\mu(A)$ on set $A \subseteq X$, generalizing

$$\int \mathbb{I}\{x \in A\} d\mu(x) = \mu(A)$$

Example 1.4 (Counting Integral)

$$\int f(x) d\#(x) = \sum_{x \in X} f(x)$$

Example 1.5 (Lebesgue Integral)

$$\int f(x) d\lambda(x) = \int \ldots \int f(x) dx_1 \ldots dx_n$$

Example 1.6 (Gaussian)

$$\int f(x)dP(x) = \int \dots \int f(x)\phi(x)dx_1 \dots dx_n = \mathbb{E}[f(Z)]$$

1.2 Densities

Given $(X, \mathcal{F})$, two measures P and μ, we say P is absolutely continuous with respect to μ if $P(A) = 0$ whenever $\mu(A) = 0$. If $P << \mu$, we can define the density function $\rho : X \to [0, \infty)$ with

$$P(A) = \int_A \rho(x)d\mu(x)$$
$$\int f(x)dP(x) = \int f(x)\rho(x)d\mu(x)$$

This is sometimes written as $\rho(x) = \frac{dP}{d\mu}(x)$ called the Radon-Nikedym derivative. If P is a proability measure, and $\mu = \lambda$, then P is called a probability density function (pdf). On the other hand if $\mu = \#$, then P is called a probability mass function (pmf).

1.3 Probability Spaces & Random Variables

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\omega \in \Omega$ called an outcome, $A \in \mathcal{F}$ called an event, and $\mathbb{P}(A)$ called the probability of A.

Definition 1 (Random Variable)**.** A random variable X is a function $X : \Omega \to \mathcal{X}$. We say X has distribution Q $(X \sim Q)$ if

$$\mathbb{P}(X \in B) = \mathbb{P}(\{\omega : X(\omega) \in B\}) = Q(B)$$

Definition 2 (Expectation)**.** We define

$$\mathbb{E}[f(X)] = \int_\Omega f(X(\omega))d\mathbb{P}(\omega) = \int_\mathcal{X} f(x)dQ(x)$$