556: MATHEMATICAL STATISTICS I

MEASURE, INTEGRATION AND PROBABILITY DISTRIBUTIONS

In the measure-theoretic framework, **random variables** are merely measurable functions with respect to the probability space $(\Omega, \mathcal{E}, \mathbb{P})$, that is, for random variable $X = X(\omega)$, $\omega \in \Omega$, with domain E, the inverse image of Borel set B, $X^{-1}(B) \equiv \{\omega \in E : X(\omega) \in B\}$, is an element of σ-algebra $\mathcal{E}$. The **expectation** of X can be written in any of the following ways

$$\int_E X(\omega)\, \mathbb{P} \qquad \int_E X(\omega)\, \mathbb{P}(d\omega) \qquad \int_E X(\omega)\, d\mathbb{P}(\omega)$$

Lebesgue-Stieltjes Integration:
If $\mathbb{P}$ is a probability measure on $\mathcal{B}$, then there is a unique corresponding real function F defined for $x \in \mathbb{R}$ by $F(x) = \mathbb{P}((-\infty, x])$, termed the **distribution function**. Conversely, if F is a distribution function, then F defines a measure μ_F on the Borel sets of $\mathbb{R}$, $\mathcal{B}$: we define μ_F on $\mathcal{B}$ via sets $(a, b])$ by

$$\mu_F((a,b]) = F(b) - F(a)$$

and then **extend** to $\mathcal{B}$ by using union operations. The probability space $(\mathbb{R}, \mathcal{B}, \mu_F)$ is then **completed** by considering and including null sets (under μ_F). Let $\mathcal{L}_F$ denote the smallest σ-algebra containing $\mathcal{B}$ and all μ_F-null sets. Thus the triple $(\mathbb{R}, \mathcal{L}_F, \mu_F)$ is the **completed** probability space.

If $g : \mathbb{R} \longrightarrow \mathbb{R}$ is a measurable function on the probability space $(\mathbb{R}, \mathcal{L}_F, \mu_F)$, then the **Lebesgue-Stieltjes integral** of g is

$$\int g\, d\mu_F = \int g(x)\, dF = \int g(x)\, dF(x).$$

The Radon-Nikodym Theorem and Change of Measure :

- *σ-finite Measure :* a measure μ defined on a σ-algebra $\mathcal{C}$ of subsets of a set C is called **finite** if $\mu(C)$ is a **finite** real number. The measure μ is called σ-finite if C the countable union of measurable sets of finite measure. A set in a measure space has σ-finite measure, if it is a union of sets with finite measure.
- *Absolute Continuity :* If μ and ν are two measures on $(C, \mathcal{C})$, then ν is **absolutely continuous** with respect to μ, denoted $\nu \ll \mu$, if and only if for all $E \in \mathcal{C}$, $\mu(E) = 0 \Longrightarrow \nu(E) = 0$.
- *The RadonNikodym Theorem :* For a measure space $(C, \mathcal{C})$, if measure ν is absolutely continuous with respect to a σ-finite measure μ, then there exists a measurable function f defined on C and taking values in $[0, \infty)$, such that

  $$\nu(A) = \int_A f\, d\mu$$

  for any measurable set A. The function f is unique almost everywhere wrt μ; it is termed the **Radon-Nikodym Derivative** of **density** of ν with respect to μ, and is often expressed as

  $$f = \frac{d\nu}{d\mu}$$

- *Change of Measure :* If μ_F is absolutely continuous wrt σ-finite measure μ, then we can rewrite

  $$\int g\, d\mu_F = \int g\, \frac{d\mu_F}{d\mu}\, d\mu = \int gf\, d\mu$$

  In practice, μ is either **counting** or **Lebesgue** measure.

Expectations :
We can construct expectations for random variables using an identical method of construction as for integral with respect to measure; for a probability space $(\Omega, \mathcal{E}, \mathbb{P})$

1 A random variable $X : \Omega \longrightarrow \mathbb{R}$ is called **simple** if it only takes finitely many distinct values; simple random variables can be written

$$X = \sum_{i=1}^{n} x_i I_{A_i}$$

for some partition $A_1, \ldots, A_n$ of Ω. The **expectation** of X is

$$\mathbb{E}[X] = \sum_{i=1}^{n} x_i \mathbb{P}(A_i)$$

2 Any non-negative random variable $X : \Omega \longrightarrow [0, \infty)$ is the **limit of some increasing sequence of simple variables**, $\{X_n\}$. That is, $X_n(\omega) \uparrow X(\omega)$, for all $\omega \in \Omega$. The expectation of X is

$$\mathbb{E}[X] = \lim_{n \longrightarrow \infty} \mathbb{E}[X_n]$$

and the limit may be infinite.

3 Any random variable $X : \Omega \longrightarrow \mathbb{R}$ can be written as $X = X^+ - X^-$, where

$$X^+(\omega) = \max\{X(\omega), 0\} \qquad X^-(\omega) = \max\{-X(\omega), 0\} = -\min\{X(\omega), 0\}$$

The expectation of X is

$$\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-]$$

if at least one of the two expectations on the right hand side is finite.

4 Thus the expectation of any random variable, $\mathbb{E}[X]$ is well-defined for every variable X such that

$$\mathbb{E}[|X|] = \mathbb{E}[X^+ + X^-] < \infty$$

Expectation defined in this fashion obeys the following rules: if $\{X_n\}$ is a sequence of rvs with $X_n(\omega) \longrightarrow X(\omega)$ for all $\omega \in \Omega$, then

(i) *Monotone Convergence:* If $X_n(\omega) \geq 0$ and $X_n(\omega) \leq X_{n+1}(\omega)$ for all n and ω, then

$$\mathbb{E}[X_n] \longrightarrow \mathbb{E}[X]$$

(ii) *Dominated Convergence:* If $|X_n(\omega)| \leq Y(\omega)$ for all n and ω, and $\mathbb{E}[|Y|] < \infty$, then

$$\mathbb{E}[X_n] \longrightarrow \mathbb{E}[X]$$

(iii) *Bounded Convergence:* If $|X_n(\omega)| \leq c$, for some c, and for all n and ω, then

$$\mathbb{E}[X_n] \longrightarrow \mathbb{E}[X]$$

These results hold even if $X_n(\omega) \longrightarrow X(\omega)$ for all ω except possibly ω in sets of probability zero (termed **null events**), that is, if $X_n(\omega) \longrightarrow X(\omega)$ **almost everywhere.**

One further result is of use in expectation calculations:

Fatou's Lemma : If $\{X_n\}$ is a sequence of rvs with $X_n(\omega) \geq Y(\omega)$ almost everywhere for all n and for some Y with $\mathbb{E}[Y] < \infty$, then

$$\mathbb{E}[\liminf_{n \longrightarrow \infty} X_n] \leq \liminf_{n \longrightarrow \infty} \mathbb{E}[X_n]$$