Measures, Integration, Convergence Note by Jon Wellner

Scribe: Mars Gao

October 22, 2022

1 Measures

Definition 1.1. [σ-algebra]. A non-void class $\mathcal{A}$ of subsets of Ω is a σ-algebra if

$$A, A_1, A_2, ... \in \mathcal{A} \Rightarrow \cup_1^\infty A_n \in \mathcal{A}.$$

Definition 1.2. A (finitely additive) measure is a function $\mu : \mathcal{A} \to [0, \infty]$ such that $\mu(\emptyset) = 0$ and $\mu(\sum A_n) = \sum \mu(A_n)$ for countable (finite) disjoint sequences A_n in $\mathcal{A}$.

Definition 1.3. A measure spase is a triple $(\Omega, \mathcal{A}, \mu)$ with $\mathcal{A}$ a σ-field and μ a measure.

Theorem 1.4. *[Caratheodory Extension Theorem] A measure μ on a field $\mathcal{C}$ can be extended to a measure on the minimal σ-field $\sigma(\mathcal{C})$ over $\mathcal{C}$. If μ is σ-finite on $\mathcal{C}$, then the extension is unique and is also σ-finite.*

Definition 1.5. A measure μ on $\mathbb{R}$ assigning finite values to finite intervals is called a Lebesgue-Stieltjes measure.

Definition 1.6. A function F on $\mathbb{R}$ which is finite, increasing, and right continuous is called a generalized distribution function

$$F(a, b] \equiv F(b) - F(a)$$

Theorem 1.7. *[Correspondence theorem] The relation*

$$\mu((a, b]) = F(a, b]$$

establishes a one-to-one correspondence between Lebesgue-Stieltjes measures μ on $\mathcal{B} = \mathcal{B}_1$ and equivalence classes of generalized df's.

2 Measurable Functions and Integration

Definition 2.1. $X : \Omega \to \mathbb{R}$ is measurable if $[x \in B] = X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$ for all $B \in \mathcal{B}_1$.

Proposition 2.2. *Suppose that $\{X_n\}$ are measurable. Then so are $\sup_n X_n, -X_n, \inf_n X_n, \liminf X_n, \limsup X_n$ and $\lim X_n$.*

Proposition 2.3. *Let X, Y be measurable. Then $X \pm Y, XY, X/Y, X^+ \equiv X\mathbf{1}_{[X\geq 0]}, |X|, g(X)$ for measurable g are all measurable.*

Proposition 2.4. *(Elementary properties) Suppose that $\int X d\mu, \int Y d\mu$ and $\int X d\mu + \int Y d\mu$ exist. Then:*
(i) $\int (X+Y) d\mu = \int X d\mu + \int Y d\mu, \int cX d\mu = c \int X d\mu$
(ii) $X \geq 0$ implies $\int X d\mu \geq 0$; $X \geq Y$ implies $\int X d\mu \geq \int Y d\mu$; and $X = Y$ a.e. implies $\int X d\mu = \int Y d\mu$.
(iii) (integrability) X is integrable if and only if $|X|$ is integrable, and either implies that X is a.e. finite. $|X| \leq Y$ with Y implies X integrable; X and Y integrable implies that $X + Y$ is integrable.

Theorem 2.5. *(Monotone convergeence theorem). If $0 \leq X_n \nearrow X$, then $\int X_n d\mu \to \int X d\mu$.*

Theorem 2.6. *(Fatou's Lemma) If $X_n \geq 0$ for all n, then $\int \underline{lim} X_n d\mu \leq \underline{lim} \int X_n d\mu$,*

Definition 2.7. A sequence X_n converges almost everywhere, denoted $X_n \to_{q.e.} X$, if $X_n(\omega) \to X(\omega)$ for all $\omega \in \Omega \backslash N$ where $\mu(N) = 0$. Note that $\{X_n\}$, X, are all defined on one measure space $(\Omega, \mathcal{A})$. If μ is a probability measure, $\mu = P$ with $P(\Omega) = 1$, we will write as a.s.

Corollary 2.8. *Let $\{X_n\}, X$ be finite measurable functions. Then $X_n \to_{a.e.} X$ if and only if*

$$\mu\left(\cap_{n=1}^{\infty} \cup_{m=n}^{\infty} |X_m - X| \geq 0\right) = 0$$

for all $\epsilon > 0$. If $\mu(\Omega) < \infty$, $X_n \to_{a.e.} X$ if and only if

$$\mu\left(\cup_{m=n}^{\infty} |X_m - X| \geq \epsilon\right) \to 0, \quad as \quad n \to \infty$$

for all $\epsilon > 0$.

Definition 2.9. (Convergence in measure; convergence in probability) A sequence of finite measurable functions X_n converge in measure to a measurable function X, denoted $X_n \to_\mu X$, if

$$\mu\left([|X_n - X| \geq \epsilon]\right) \to 0$$

for all $\epsilon > 0$. If μ is a probability measure, $\mu(\Omega) = 1$, call $\mu = P$, write $X_n \to_\mu X$, and say X_n converge in probability to X.

Proposition 2.10. *Let X_n's be finite a.e.*
(i) If $X_n \to_\mu X$ then there exist a subsequence $\{n_k\}$ such that $X_{n_k} \to_{a.e.} X$.
(ii) If $\mu(\Omega) < \infty$ and $X_n \to_{a.e.} X$, then $X_n \to_\mu X$.

Theorem 2.11. *(Dominated Convergence Theorem) If $|X_n| \leq Y$ a.e. with Y integrable, and if $X_n \to_\mu X$ (or $X_n \to_{a.e.} X$), then $\int |X_n - X| \, d\mu \to 0$ and $\lim \int X_n d\mu = \int X d\mu$.*

Definition 2.12. Let X be a finite measurable function on probability space $(\Omega, \mathcal{A}, P)$ (so that $P(\Omega) = 1$). Then X is called a random variable and

$$P_X(B) \equiv P(X \in B) = P\left(\{\omega \in \Omega : X(\omega) \in B\}\right)$$

for all $B \in \mathcal{B}$ is called the induced probability distribution of X (on $\mathbb{R}$). Thus $(\mathbb{R}, \mathcal{B}, P_X)$ is a probability space.

Theorem 2.13. *(Theorem of the unconscious statistician.) If g is a finite measurable function from $\mathbb{R}$ to $\mathbb{R}$, then*

$$\int_\Omega g(X(\omega)) dP(\omega) = \int_{\mathbb{R}} g(x) dP_X(x) = \int_{\mathbb{R}} g(x) dF_X(x).$$

Proposition 2.14. *(Interchange of integral and limit or derivative.) Suppose that $X(\omega, t)$ is measurable for each $t \in (a, b)$.*
(i) If $X(\omega, t)$ is a.e. continuous in t at t_0 and $|X(\omega, t)| \leq Y(\omega)$ a.e. for $|t - t_0| < \delta$ with Y integrable, then $\int X(\cdot, t) d\mu$ is continuous in t at t_0.
(ii) Suppose that $\frac{\partial}{\partial t} X(\omega, t)$ exists for a.e. ω, all $t \in (a, b)$, and $\left|\frac{\partial}{\partial t} X(\omega, t)\right| \leq Y(\omega)$ integrable a.e. for all $t \in (a, b)$

$$\frac{\partial}{\partial t} \int_\Omega X(\omega, t) d\mu(\omega) = \int_\Omega \frac{\partial}{\partial t} X(\omega, t) d\mu(\omega)$$

3 Absolute Continuity, Radon-Nikodym Theorem, Fubini's Theorem

Definition 3.1. The measure ν defined by

$$\nu(A) \equiv \int_A X d\mu = \int_\Omega \mathbf{1}_A X d\mu$$

is said to have density X w.r.t. μ.

Definition 3.2. If μ, ν are two measures on $(\Omega, \mathcal{A})$ s.t. $\mu(A) = 0$ implies $\nu(A) = 0$ for any $A \in \mathcal{A}$, then ν is said to be absolutely continuous w.r.t. μ, and we could write $\nu \ll \mu$. We also say that ν is dominated by μ.

Theorem 3.3. *(Radon-Nikodym theorem.) Let $(\Omega, \mathcal{A}, \mu)$ be a σ-finite measure space, and let ν be a measure on $(\Omega, \mathcal{A})$ with $\nu \ll \mu$. Then there exists a measurable function $X \geq 0$ such that $\nu(A) = \int_A X d\mu$ for all $A \in \mathcal{A}$. The function $X \equiv \frac{d\nu}{d\mu}$ is unique in the sense that if Y is another such function, then $Y = X$ a.e. w.r.t. μ. X is called the Radon-Nikodym derivative of ν w.r.t. μ.*

Corollary 3.4. *(Change of Variable Theorem.) Suppose that ν, μ are σ-dinite measures defined on a measure space $(\Omega, \mathcal{A})$ with $\nu \ll \mu$, and suppose that Z is a measurable function such that $\int Z d\nu$ is well-defined. Then for all $A \in \mathcal{A}$,*

$$\int_A Z d\nu = \int_A Z \frac{d\nu}{d\mu} d\mu$$

Proposition 3.5. *(Scheffe's theorem.) Suppose that $\nu_n(A) = \int_A f_n d\mu$, that $\nu(A) = \int_A f d\nu$ where f_n are densities and $\nu_n(\Omega) = \nu(\Omega) < \infty$ for all n, and that $f_n \to f$ a.e. μ. Then*

$$\sup_{A \in \mathcal{A}} |\nu_n(A) - \nu(A)| = \frac{1}{2} \int_\Omega |f_n - f| \to 0.$$

Theorem 3.6. *(Fubini-Tonelli theorem.) Suppose that $f : \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$ is $\mathcal{X} \times \mathcal{Y}$-measurable and $f \geq 0$. Then*

$$\int_{\mathbb{Y}} f(x,y) d\nu(y) \quad is \quad \mathcal{X} - measurable$$

$$\int_{\mathbb{Y}} f(x,y) d\nu(y) \quad is \quad \mathcal{X} - measurable$$

and

$$\int_{\mathbb{X} \times \mathbb{Y}} f(x,y) d\pi(x,y) = \int_{\mathbb{X}} \left\{ \int_{\mathbb{Y}} f(x,y) d\nu(y) \right\} d\mu(x) = \int_{\mathbb{Y}} \left\{ \int_{\mathbb{X}} f(x,y) d\mu(x) \right\} d\nu(y).$$

If $f \in L_1(\pi)$ (so $\int_{\mathbb{X} \times \mathbb{Y}} |f| \, d\pi < \infty$), then the above equation holds.