Math 7351    1. Measures  

Spring 2005

Definition A ring on \( X \) is a non-empty collection \( A \) of sets such that \( A,B \in A \Rightarrow A \setminus B \in A \) and \( A \cup B \in A \). It is a \( \sigma \)-ring if \( A_1,A_2,\ldots \in A \Rightarrow \bigcup_{i=1}^{\infty} A_i \in A \). An algebra (\( \sigma \)-algebra) is a ring (\( \sigma \)-ring) containing the set \( X \).

For algebras one can replace the condition \( A \setminus B \in A \) by \( X \setminus B \in A \).

Both (\( \sigma \))-rings and (\( \sigma \))-algebras are also closed under finite (countable) intersections.

Definition A measurable space is a pair \((X,A)\) where \( A \) is a \( \sigma \)-algebra on \( X \).

Definition A measure \( \mu \) on \((X,A)\) is a function \( \mu : A \rightarrow [0,\infty] \) that is countably additive: If \( A_i \in A \) are disjoint sets for \( i \in I \), and \( I \) is countable, then \( \mu(\bigcup_{i \in I} A_i) = \sum_{i \in I} \mu(A_i) \).

[Note: we include finite \( I \) and empty \( I \), so in particular \( \mu(\emptyset) = 0 \).]

Definition We say \( \mu \) is finite if \( \mu(X) < \infty \). We say \( \mu \) is \( \sigma \)-finite if \( X = \bigcup_{i=1}^{\infty} X_i \) with \( \mu(X_i) < \infty \). We call \( \mu \) a probability measure if \( \mu(X) = 1 \).

Definition A measure space is a triple \((X,A,\mu)\) where \( A \) is a \( \sigma \)-algebra on \( X \) and \( \mu \) is a measure on \((X,A)\). We say \( A \subseteq X \) is \( \mu \)-measurable if \( A \in A \).

Examples

1. If \( \mathcal{L} \) is the set of Lebesgue measurable sets and \( \lambda \) is the Lebesgue measure, then \((\mathbb{R},\mathcal{L},\lambda)\) is a (\( \sigma \)-finite) measure space. More generally, if \( f \geq 0 \) is measurable and \( \mu(S) = \int_S f(x) \, dx \) then \( \mu \) is a measure on \((\mathbb{R},\mathcal{L})\).

2. If \( X \) is any set, the counting measure \( \mu(A) = |A| \) is a measure on \((X,\mathcal{P}(X))\). It is finite (\( \sigma \)-finite) iff \( X \) is finite (countable). More generally, if \( w : X \rightarrow [0,\infty] \) is any function, then the weighted counting measure \( \mu(A) = \sum_{x \in A} w(x) \) is a measure on \((X,\mathcal{P}(X))\).

Lemma 1. Suppose \((X,A,\mu)\) is a measure space. Then

1. \( \mu \) is monotonic: if \( A \subseteq B \) then \( \mu(A) \leq \mu(B) \).

2. \( \mu \) is countably subadditive: if \( A_i \in A \), \( I \) countable, then \( \mu(\bigcup_{i \in I} A_i) \leq \sum_{i \in I} \mu(A_i) \).

3. If \( A_1 \subseteq A_2 \subseteq \ldots \), then \( \mu(\bigcup_{i=1}^{\infty} A_i) = \lim_{i \to \infty} \mu(A_i) \).

4. If \( A_1 \supseteq A_2 \supseteq \ldots \) and \( \mu(A_1) < \infty \), then \( \mu(\bigcap_{i=1}^{\infty} A_i) = \lim_{i \to \infty} \mu(A_i) \).

Definition \((X,A,\mu)\) is complete if \( \mu(A) = 0 \) implies all subsets of \( A \) lie in \( A \).

Lemma 2. If \((X,A,\mu)\) is a measure space, then there is a unique complete measure space \((X,\mathcal{A},\hat{\mu})\) with \( \mathcal{A} = \{ A \cup E : A \in A, E \subseteq B \in A, \mu(B) = 0 \} \) and \( \hat{\mu}|_{\mathcal{A}} = \mu \).

The space \((X,\mathcal{A},\hat{\mu})\) is called the completion of \((X,A,\mu)\).
Math 7351 2. Signed Measures Spring 2005

**Definition** Given a measurable space \((X, \mathcal{A})\), a *signed measure* is a countably additive function \(\mu: \mathcal{A} \to \mathbb{R}\) such that either \(\mu(A)\) is never \(+\infty\) or it is never \(-\infty\). We call \(\mu\) finite if \(\mu(A)\) is never \(\pm\infty\).

The conditions on \(\pm\infty\) imply we never get \(-\infty - \infty\) in the ‘countably additive’ property.

**Definition** A set \(A \in \mathcal{A}\) is positive if \(\mu(B) \geq 0\) for all \(B \subseteq A\), negative if \(\mu(B) \leq 0\) for all \(B \subseteq A\), and null if \(\mu(B) = 0\) for all \(B \subseteq A\), \(B \in \mathcal{A}\).

**Theorem (Hahn decomposition)** If \(\mu\) is a signed measure, then any \(A \in \mathcal{A}\) can be written as disjoint union \(A = A^+ \cup A^-\) where \(A^+\) is positive and \(A^-\) is negative.

**Proof.** W.l.o.g., assume \(\mu\) is never \(+\infty\). Pick any \(B_0 \subseteq A\) with \(\mu(B_0) \neq -\infty\). If there is a \(C_0 \subseteq B_0\) with \(\mu(C_0) < 0\), pick \(C_0\) with \(\mu(C_0) < \frac{1}{2}\inf\{\mu(C) : C \subseteq B_0\}\) (\(< -1\) if \(\inf = -\infty\)) and let \(B_1 = B_0 \setminus C_0\). Repeat this process to get a sequence \(B_0 \supseteq B_1 \supseteq B_2 \supseteq \ldots\) and let \(B = \bigcap B_n\). Then \(\mu(B_0 \setminus B) = \sum \mu(C_i) < 0\), so \(\mu(B) \geq \mu(B_0)\). By assumption \(\mu(B) < \infty\), so \(\mu(C_i) \to 0\). Thus if \(C \subseteq B\) and \(\mu(C) < 0\) then some \(\mu(C_i) > \frac{1}{2}\mu(C)\), contradicting the choice of \(C_i\). Thus \(B\) is positive and \(\sup\{\mu(B) : B \subseteq A\} = \mu(B) : B \subseteq A, B \text{ positive}\). Thus we can find a sequence of positive sets \(B_i\) with \(\mu(B_i) \to \sup\{\mu(B) : B \subseteq A\}\). Let \(A^+ = \bigcup B_i\). If \(C \subseteq A^+\) then \(C = \bigcup(B_i \cap C \setminus \bigcup_{j<i} B_j)\) is a disjoint union of subsets of the \(B_i\), so \(\mu(C) \geq 0\). Thus \(A^+\) is positive and \(\mu(A^+) = \mu(B_i) + \mu(A^+ \setminus B_i) \geq \mu(B_i)\) for all \(i\), so \(\mu(A^+) = \sup\{\mu(B) : B \subseteq A\}\). Let \(A^- = A \setminus A^+\). If \(C \subseteq A^-\) with \(\mu(C) > 0\) then \(\mu(A^+ \cup C) > \mu(A^+)\), a contradiction. Hence \(A^-\) is negative. \(\square\)

Note: The decomposition \(A = A^+ \cup A^-\) is not unique in general.

**Definition** A (signed) measure \(\mu\) is supported on a subset \(A \in \mathcal{A}\) if \(\mu(B) = \mu(B \cap A)\) for all \(B \in \mathcal{A}\). Equivalently, \(\mu(B) = 0\) for all \(B \subseteq A^c\). Two (signed) measures \(\mu\) and \(\nu\) are mutually singular, \(\mu \perp \nu\), if they are supported on disjoint sets.

**Theorem (Jordan decomposition)** If \(\mu\) is a signed measure then \(\mu = \mu^+ - \mu^-\) where \(\mu^\pm\) are mutually singular measures, at least one of which is finite. Moreover, this decomposition is unique.

**Proof.** Write \(X = X^+ \cup X^-\) as above and set \(\mu^+(A) = \mu(A \cap X^+)\) and \(\mu^-(A) = -\mu(A \cap X^-)\). Then \(\mu = \mu^+ - \mu^-\) and \(\mu^\pm\) are mutually singular measures. Assume now that \(\mu = \mu^+ - \mu^- = \nu^+ - \nu^-\) and \(X = Y^+ \cup Y^-\) with \(\nu^\pm\) supported on \(Y^\pm\). Now if \(A \subseteq X^+ \cap Y^-\), \(\mu(A) = \mu^+(A) = -\nu^-(A)\), so \(\mu^+(A) = -\nu^-(A) = 0\). Hence if \(A \subseteq X^+\) then \(\nu^-(A) = 0\) and \(\mu(A) = \mu^+(A) = \nu^+(A)\). Similarly if \(A \subseteq X^-\) then \(\nu^+(A) = 0\), so for any \(A\), \(\mu^+(A) = \nu^+(A)\). Thus \(\mu^+ = \nu^+\), so by subtraction, \(\mu^- = \nu^-\). \(\square\)

**Exercise:** Suppose \(f\) is integrable. Show that \(\mu(S) = \int_S f(x) \, dx\) is a signed measure. Give an expression for \(\mu^\pm(S)\).
Math 7351  3. Constructing Measures   Spring 2005

Definition  A semiring on $X$ is a non-empty collection $\mathcal{I}$ of subsets of $X$ such that

S1. $I, J \in \mathcal{I} \implies I \cap J \in \mathcal{I}$,
S2. $I, J \in \mathcal{I} \implies I \setminus J$ is a finite disjoint union of elements of $\mathcal{I}$.

A semialgebra is a semiring containing $X$.

Examples  
1. The set of all half-open intervals $(a, b)$, $a, b \in \mathbb{R}$.
2. The set of rectangles $A \times B$ in $X \times Y$.

Lemma 1.  Let $\mathcal{I}$ be a semiring.

1. If $A_1, \ldots, A_n \in \mathcal{I}$, then $\exists$ disjoint $I_1, \ldots, I_N$ with each $A_i$ a union of some $I_j$'s.
2. Any element of the ring generated by $\mathcal{I}$ is a finite disjoint union of elements of $\mathcal{I}$,
3. Any countable union of elements of $\mathcal{I}$ is a disjoint countable union of elements of $\mathcal{I}$.

Proof.  1. Induction: replace $I_i$ with $I_i \cap A_{n+1}$ and the disjoint sets with union $I_i \setminus A_{n+1}$.
   By induction on $N$ one can also decompose $A_{n+1} \setminus \bigcup_{i=1}^N I_i$ as a disjoint union.
2. Clear. 3. Write $\bigcup A_i$ as a disjoint union of $A_i \setminus \bigcup_{j<i} A_j$, each of which is a finite disjoint union of elements of $\mathcal{I}$.

We say a function $l: \mathcal{I} \to [0, \infty]$ is a measure on $\mathcal{I}$ if it is countably additive when defined:
if $I_i \in \mathcal{I}$, are disjoint, $I$ is countable, and $\bigcup_{i \in I} I_i \in \mathcal{I}$, then $l(\bigcup_{i \in I} I_i) = \sum_{i \in I} l(I_i)$.

We shall prove:

Theorem (Carathéodory)  Suppose $l$ is a measure on the semiring $\mathcal{I}$. Then there is an extension of $l$ to a measure $\mu$ on some $\sigma$-algebra containing $\mathcal{I}$. Moreover, this measure is uniquely determined on the $\sigma$-ring generated by $\mathcal{I}_{\text{fin}} = \{ I \in \mathcal{I} : l(I) < \infty \}$.

Definition  Suppose $\mathcal{I}$ is any collection of subsets of $X$ and $l: \mathcal{I} \to [0, \infty]$ any function. Define for any $A \subseteq X$, $\mu^*(A) = \inf_{A \subseteq \bigcup I_i} \sum l(I_i)$, where the infimum is over all countable collections of $I_i \in \mathcal{I}$ with $A \subseteq \bigcup I_i$.

We include finite and empty collections, so in particular $\mu^*(\emptyset) = 0$.
Also, if there is no countable collection of $I_i$ with $A \subseteq \bigcup I_i$ then $\mu^*(A) = \infty$.

Lemma 2.  For any $l: \mathcal{I} \to [0, \infty]$, $\mu^*$ is an outer measure, i.e.,

1. $\mu^*$ is monotonic: if $A \subseteq B$ then $\mu^*(A) \leq \mu^*(B)$,
2. $\mu^*$ is countably subadditive: if $\{ A_i : i \in I \}$ is countable, $\mu^*(\bigcup_{i \in I} A_i) \leq \sum_{i \in I} \mu^*(A_i)$.

Definition  If $\mu^*$ is an outer measure, we say $A \subseteq X$ is $\mu^*$-measurable if for all $E \subseteq X$, $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \setminus A)$. [Subadditivity $\Rightarrow \leq$, so we only need $\geq$.]
Lemma 3. The set $\mathcal{A}$ of all $\mu^*$-measurable sets is a $\sigma$-algebra and the restriction of $\mu^*$ to $\mathcal{A}$ is a complete measure.

Proof. Clearly $A = X$ is measurable and $A$ is measurable iff $X \setminus A$ is measurable. Suppose $A_1, A_2, \ldots$ are measurable and let $A = \bigcup A_i$. Define inductively $E_0 = E$ and $E_{i+1} = E_i \setminus A_i$. By measurability of $A_i$, $\mu^*(E_i) = \mu^*(E_i \cap A_i) + \mu^*(E_{i+1})$. Hence

$$\mu^*(E) = \sum_{i=1}^{\infty} \mu^*(E_i \cap A_i) + \mu^*(E_{n+1}).$$

However $E \setminus A \subseteq E_{n+1}$, so $\mu^*(E) \geq \sum_{i=1}^{n} \mu^*(E_i \cap A_i) + \mu^*(E \setminus A)$ for all $n$. Thus

$$\mu^*(E) \geq \sum_{i=1}^{\infty} \mu^*(E_i \cap A_i) + \mu^*(E \setminus A). \quad (1)$$

However, $\bigcup (E_i \cap A_i) = E \cap A$, so by subadditivity, $\mu^*(E \cap A) \leq \sum_{i=1}^{\infty} \mu^*(E_i \cap A_i)$. Thus

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \setminus A),$$

as required. (If there are only finitely many $A_i$, set the other $A_i = \emptyset$.)

If $A_i$ are disjoint and $\mu^*$-measurable, take $E = A$ so that $E_i \cap A_j = A_i$ and (1) gives $\mu^*(A) \geq \sum_{i=1}^{\infty} \mu^*(A_i)$. Since $\mu^*$ is countably subadditive, $\mu^*(A) = \sum_{i=1}^{\infty} \mu^*(A_i)$.

For completeness, note that if $\mu^*(A) = 0$ and $B \subseteq A$ then $\mu^*(E \cap B) \leq \mu^*(A) = 0$ and $\mu^*(E \setminus B) \leq \mu^*(E)$, so $\mu^*(E) \geq \mu^*(E \cap B) + \mu^*(E \setminus B)$ and so $B$ is $\mu^*$-measurable. □

Lemma 4. If $\mathcal{I}$ is a semiring and $l$ is a measure on $\mathcal{I}$ then every $I \in \mathcal{I}$ is $\mu^*$-measurable and $\mu^*(I) = l(I)$.

Proof. Fix $I \in \mathcal{I}$. Assume $E \subseteq \bigcup I_i$, and $\mu^*(E) \geq \sum l(I_i) - \varepsilon$. Now $\mu^*(E \cap I) \leq \sum l(I_i \cap I)$, and $\mu^*(E \setminus I) \leq \sum l(I_i \setminus I)$ where $I_i \setminus I = \bigcup I_i \setminus I_j$ is a disjoint union. But by assumption $l(I_i \setminus I) = l(I_i \cap I) + \sum l(I_i \cap I_j)$, so $\mu^*(E) \geq \mu^*(E \cap I) + \mu^*(E \setminus I) - \varepsilon$. Since this is true for all $\varepsilon > 0$, $I$ is $\mu^*$-measurable. Clearly $\mu^*(I) \leq l(I)$. Suppose $I \subseteq \bigcup I_i$. Let $J_i = I \cap I_i \setminus \bigcup_{j<i} I_j$. By Lemma 1, both $J_i$ and $I_i \setminus J_i$ are finite disjoint unions of elements of $\mathcal{I}$, $J_i = \bigcup I_i \setminus J_i$, and $I_i \setminus J_i = \bigcup J_i \setminus J_i$. But $I$ is a disjoint union of the $J_i$, so $l(I) = \sum l(J_i)$. Now $l(I_i) = \sum l(I_i \setminus J_i) + \sum l(J_i)$, so $\sum l(I_i) \geq l(I)$ and thus $\mu^*(I) = l(I)$. □

Lemma 5. If $\mathcal{I}$ is a semiring and $l$ is a measure on $\mathcal{I}$ then any extension of $l$ to a measure $\nu$ on a $\sigma$-algebra containing $\mathcal{I}$ satisfies $\nu \leq \mu^*$. Moreover, $\nu = \mu^*$ on the $\sigma$-ring generated by $\mathcal{I}_{\text{fin}}$.

Proof. Let $A$ be $\nu$-measurable. If $A \subseteq \bigcup I_i$ then $\nu(A) \leq \sum \nu(I_i) = \sum l(I_i)$, so $\nu(A) \leq \mu^*(A)$. Now assume $A$ is in the $\sigma$-ring generated by $\mathcal{I}_{\text{fin}}$. Then $A \subseteq \bigcup I_i$ for some $I_i \in \mathcal{I}_{\text{fin}}$. (The collection of all such $A$ is a $\sigma$-ring and contains $\mathcal{I}_{\text{fin}}$.) Thus by Lemma 1, $A \subseteq \bigcup I_i$ for some disjoint $I_i \in \mathcal{I}_{\text{fin}}$. Now $\nu(I_i \setminus A) + \nu(I_i \cap A) = \nu(I_i) = \mu^*(I_i) = \mu^*(I_i \setminus A) + \mu^*(I_i \cap A)$, and $\nu \leq \mu^*$, so $\nu(I_i \cap A) = \mu^*(I_i \cap A)$ and $\nu(A) = \sum \nu(I_i \cap A) = \sum \mu^*(I_i \cap A) = \mu^*(A)$. □

Example Suppose $\mu(A) = |A|$ and $\nu(A) = 2|A|$ for $A \subseteq \mathbb{R}$. Let $\mathcal{I} = \{(a, b] : a, b \in \mathbb{R}\}$. Then $\mu|_{\mathcal{I}} = \nu|_{\mathcal{I}}$ but $\mu \neq \nu$ on singletons, which are in the $\sigma$-ring generated by $\mathcal{I}$.

The Carathéodory Theorem follows from Lemmas 3–5.
Let $\mathcal{B}$ be the Borel sets of $\mathbb{R}$. If $\mu$ is a finite measure on $(\mathbb{R}, \mathcal{B})$, then the cumulative distribution function of $\mu$ is

$$F(x) = \mu((−\infty,x]).$$

Note that $\mu((a,b]) = F(b) − F(a)$ for all $a \leq b$ and $F$ is an increasing function of $x$ that is continuous on the right:

$$F(a) \leq \lim_{x \to a^+} F(x) \leq \lim_n F(a + \frac{1}{n}) = \mu(\bigcap_n (−\infty,a + \frac{1}{n}]) = \mu((−\infty,a]) = F(a).$$

**Theorem** If $F$ is an increasing real valued function that is continuous on the right, then there is a unique measure $\mu_F$ on $(\mathbb{R}, \mathcal{B})$ with $\mu_F((a,b]) = F(b) − F(a)$ for all $a \leq b$.

**Proof.** Let $\mathcal{I} = \{(a,b) : a \leq b\}$. Then $\mathcal{I}$ is a semiring. Define $l : \mathcal{I} \to [0,\infty]$ by $l((a,b]) = F(b) − F(a)$. We shall show that $l$ is a measure on $\mathcal{I}$.

Suppose $(a,b] = \bigcup_{i=1}^{\infty} (a_i,b_i]$ is a disjoint union. For any $N$ one can define $(c_j,d_j]$, $j \leq N$, to be $(a_i,b_i]$, $i \leq N$, ordered in increasing order of $a_i$. Set $d_0 = a$ and $c_{N+1} = b$. Then

$$a = d_0 \leq c_1 \leq d_1 \leq c_2 \leq \cdots \leq d_N \leq c_{N+1} = b,$$

$$F(b) − F(a) = \sum_{i=1}^{N}(F(d_i) − F(c_i)) + \sum_{i=0}^{N}(F(c_{i+1}) − F(d_i)) \geq \sum_{i=1}^{N}(F(b_i) − F(a_i)),$$

since $F$ increasing. Thus $l((a,b]) \geq \sum_{i=1}^{N} l((a_i,b_i])$ for each $N$, so $l((a,b]) \geq \sum_{i=1}^{\infty} l((a_i,b_i])$.

Fix $\varepsilon > 0$. Then there is a $\delta$ with $F(a+\delta) < F(a) + \varepsilon$ and $\delta_i$ with $F(b_i+\delta_i) < F(b_i) + \varepsilon/2^i$. The open sets $(a_i,b_i+\delta_i]$ cover the compact set $[a+\delta,b]$. Hence there is a finite collection of sets $(a_i,b_i+\delta_i]$ that cover $(a+\delta,b]$. Inductively removing any $(a_i,b_i+\delta_i]$ that lie in some other $(a_j,b_j+\delta_j]$ and ordering the remaining sets by $a_i$, we obtain intervals $(c_i,d_i]$ with $c_{i+1} \leq d_i$. Setting $d_0 = a + \delta$, $c_{N+1} = b$, we may assume this also holds with $i = 0, N$.

Since $F$ is increasing

$$F(b) − F(a + \delta) = \sum_{i=1}^{N}(F(d_i) − F(c_i)) − \sum_{i=0}^{N}(F(d_i) − F(c_{i+1})) \leq \sum(F(b_i + \delta_i) − F(a_i)).$$

Thus

$$F(b) − F(a) \leq \sum(F(b_i) − F(a_i)) + \varepsilon + \sum \varepsilon/2^i.$$

So $l((a,b]) \leq \sum_{i=1}^{\infty} l((a_i,b_i]) + 2\varepsilon$ for any $\varepsilon > 0$. Hence $l$ is a measure on $\mathcal{I}$.

Finally, the $\sigma$-ring generated by $\mathcal{I}_{\text{fin}} = \mathcal{I}$ contains all open intervals since $(a,b) = \bigcup(a_i,b_i]$ when $a_i$ decreases to $a \in (−\infty,\infty)$ and $b_i$ increases to $b \in (−\infty,\infty)$. Thus it contains all open sets (each is a countable union of open intervals), and so all Borel sets. The result now follows from Carathéodory.

**Examples**

1. Lebesgue measure can be constructed as the special case $F(x) = x$.
2. Let $F$ be the Cantor Ternary function. Then $\mu_F$ is supported on a set of Lebesgue measure zero (the Cantor set), but is zero on all singletons.

One can extend this result to (finite) signed measures, if we replace the condition that $F$ is increasing by the condition that $F$ is has bounded variation, since in this case one can write $F = G − H$ where $G$ and $H$ are (bounded) increasing functions and define $\mu_F = \mu_G − \mu_H$. 

---

**Math 7351 4. Lebesgue-Stieltjes Spring 2005**

---
Math 7351  5. Product Measures   Spring 2005

**Theorem (Weak Monotone Convergence Theorem)** Suppose \((X, \mathcal{A}, \mu)\) is a measure space and \(A, A_i \in \mathcal{A}, c, c_i \geq 0\). If \(c \mu(A) > \sum_{i=1}^{\infty} c_i \mu(A_i)\) then \(\exists x \in A: c > \sum_{i} x \in A_i c_i\).

**Proof.** Pick \(\gamma < c\) and \(\alpha < \mu(A)\) so that \(\gamma \alpha > \sum_{i=1}^{\infty} c_i \mu(A_i)\). Let \(S_n = \{x \in A : \sum_{i \leq n, x \in A_i} c_i \geq \gamma\}\). Then \(S_n\) is a union of intersections of the sets \(A_1, \ldots, A_n\), so is measurable. If \(\bigcup S_n = A\) then \(\mu(S_n) \to \mu(A)\), so \(\exists N: \mu(S_N) > \alpha\). Let \(I_1, \ldots, I_M\) be disjoint elements of \(\mathcal{A}\) such that each \(A_i, i \leq N\), and \(A\) can be written as a union of some of the \(I_s\).

Then \(S_N\) is a disjoint union of some of the \(I_s\) and

\[
\sum_{i=1}^{\infty} c_i \mu(A_i) \geq \sum_{i=1}^{N} \sum_{I_s \subseteq A_i} c_i \mu(I_s) = \sum_{I_s} \sum_{i \leq N, I_s \subseteq A_i} c_i \mu(I_s) \geq \sum_{I_s \subseteq S_N} \gamma \mu(I_s) > \gamma \alpha,
\]

a contradiction. Hence \(\bigcup S_n \neq A\) and there is an \(x \in A\) with \(c > \gamma \geq \sum_{i} x \in A_i c_i\). \(\square\)

**Theorem** If \((X, \mathcal{A}, \mu)\) and \((Y, B, \nu)\) are measure spaces then there is a measure \(\mu \times \nu\) on the \(\sigma\)-algebra \(\mathcal{A} \otimes \mathcal{B}\) generated by \(\mathcal{A} \times \mathcal{B}\) with \((\mu \times \nu)(A \times B) = \mu(A)\nu(B)\) for all \(A \in \mathcal{A}\), \(B \in \mathcal{B}\). Moreover, if \(\mu\) and \(\nu\) are both \(\sigma\)-finite then this measure is unique and \(\sigma\)-finite.

**Proof.** Let \(\mathcal{I} = \{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}\) and let \(l(A \times B) = \mu(A)\nu(B)\). Now \((A \times B) \cap (A' \times B') = (A \cap A') \times (B \cap B')\) and \((A \times B) \setminus (A' \times B')\) is the disjoint union of \((A \setminus A') \times B\) and \(A' \times (B \setminus B')\). Hence the measurable rectangles form a semiring.

Suppose \(A \times B = \bigcup_{i=1}^{\infty} (A_i \times B_i)\) is a disjoint union. Let \(c = \nu(B), c_i = \nu(B_i)\), then for all \(x, B = \bigcup_{i: x \in A_i} B_i\), so \(c \leq \sum_{i: x \in A_i} c_i\). By WMCT, \(l(A \times B) = c \mu(A) \leq \sum c_i \mu(A_i) = \sum l(A_i \times B_i)\). Now fix \(N\) and construct disjoint \(I_1, \ldots, I_M\) so that each \(A_i, i \leq N\), is a union of some of the \(I_s\).

\[
\sum_{i=1}^{N} \mu(A_i) \nu(B_i) \leq \sum_{i \leq N} \sum_{I_s \subseteq A_i} \mu(I_s) \nu(B_i) \leq \sum_{I_s \subseteq A} \mu(I_s) \sum_{i \leq N, I_s \subseteq A_i} \nu(B_i) \leq \sum_{I_s \subseteq A} \mu(I_s) \nu(B) \leq \mu(A) \nu(B).
\]

Letting \(N \to \infty\) gives \(\sum l(A_i \times B_i) \leq l(A \times B)\). Thus \(l\) is a measure and the result follows from Carathéodory. \(\square\)

Define \(\mu \cdot \nu\) to be the completion of \(\mu \times \nu\), with \(\sigma\)-algebra \(\mathcal{A} \otimes \mathcal{B}\).

If \(E \subseteq X \times Y\), define the section of \(E\) at \(x\) to be \(E_x = \{y : (x, y) \in E\}\).

We say a property holds \(\mu\)-a.e. if the set of points where it fails has \(\mu\)-measure zero.

**Lemma** If \(E\) is \(\mu \times \nu\)-measurable, then \(E_x\) is \(\nu\)-measurable for all \(x \in X\).

**Proof.** The set \(\{E \subseteq X \times Y : E_x \text{ is } \nu\text{-measurable for all } x\}\) is a \(\sigma\)-algebra and contains all measurable rectangles \(A \times B\), so contains \(\mathcal{A} \otimes \mathcal{B}\). \(\square\)

Note that this is not true for \(\mu \cdot \nu\) measurable sets. E.g., if \(S\) is a non Lebesgue measurable set in \(\mathbb{R}\) then \(E = \{x\} \times S \subseteq \{x\} \times \mathbb{R}\) is a subset of a set of measure zero, so is \(\lambda \lambda\)-measurable, but \(E_x\) is not measurable.
Suppose \((X_i, A_i, \mu_i), i = 1, 2, \ldots\) are measure spaces with \(\mu_i(X_i) = 1\), we shall construct a measure on \(X = \prod X_i\).

**Definition** A cylinder set is a set of the form \(A = \prod A_i\) where \(A_i \in A_i\) and \(A_i = X_i\) for all but finitely many \(i\).

**Theorem** There exists a unique probability measure on the \(\sigma\)-algebra generated by cylinder sets of \(X = \prod X_i\) in which each cylinder set \(\prod A_i\) gets measure \(\prod \mu_i(A_i)\).

Note: \(\prod \mu_i(A_i)\) is really a finite product since \(\mu_i(A_i) = 1\) for all but finitely many \(i\)’s.

**Proof.** For each \(N\) and each cylinder set \(A = \prod A_i\), define \(A^{(N)} = \prod_{i > N} A_i\) and \(A^{(N)} = \prod_{i \leq N} A_i\), so that one can regard \(A\) as a product \(A^{(N)} \times A^{(N)}\). Since \(A\) is a cylinder set, \(A^{(N)} = X^{(N)}\) for sufficiently large \(N\). Define \(l(A) = \prod \mu_i(A_i)\), and more generally \(l(A^{(N)}) = \prod_{i > N} \mu_i(A_i)\). By the existence of finite product measures, there are measures \(\mu(N)\) on \(X\) with \(\mu(N)(A) = l(A)\) for all cylinder sets with \(A^{(N)} = X^{(N)}\).

Suppose \(A\) and \(A_i\) are cylinder sets with \(A\) a disjoint union of the \(A_i\). Now \(A \supseteq \bigcup_{i=1}^n A_i\), and for sufficiently large \(N\), \(A^{(N)} = A_1^{(N)} = \cdots = A_n^{(N)} = X^{(N)}\). Thus \(l(A) = \mu(N)(A) \geq \sum_{i=1}^n \mu(N)(A_i) = \sum_{i=1}^n l(A_i)\). Letting \(n \to \infty\), \(l(A) \geq \sum_{i=1}^\infty l(A_i)\).

Suppose \(l(A) > \sum_{i=1}^\infty l(A_i)\). We shall construct a point \(x = (x_1, x_2, \ldots) \in A\) that is not in any \(A_i\). Assume we have defined \(x_1, \ldots, x_{N-1}\) and let \(X_{N-1} = \{x_1\} \times \cdots \times \{x_{N-1}\} \times X^{(N-1)}\) be the set of all points in \(X\) with first \(N - 1\) components equal to \(x_i\). Assume that \(X_{N-1} \cap A \neq \emptyset\) and

\[
l(A^{(N-1)}) > \sum_{i : x_{N-1} \cap A_i \neq \emptyset} l(A_i^{(N-1)}).
\]

Since \(X_0 = X\), this holds for \(N = 1\). Write \(c = l(A^{(N)})\) and \(c_i = l(A_i^{(N)})\). Then \(l(A^{(N-1)}) = c \mu(N)((A)_{N})\) and \(l(A_i^{(N-1)}) = c_i \mu(N)(A_i_{N})\). Thus by the WMCT there exists an \(x_N \in (A)_{N}\) (so \(X_N \cap A \neq \emptyset\) with

\[
l(A^{(N)}) = c > \sum_{i : x_N \cap A_i \neq \emptyset, x_N \in (A)_{N}} c_i = \sum_{i : x_N \cap A_i \neq \emptyset} l(A^{(N)}).
\]

Now fix \(i\). If \((x_1, \ldots) \in A_i\) then for sufficiently large \(N\), \(l(A_i^{(N)}) = 1 > l(A^{(N)})\), a contradiction. But for large enough \(N\), \(A^{(N)} = X^{(N)}\), so \((x_1, \ldots) \in X_{N} \subseteq A\). Thus \(A \neq \bigcup A_i\), a contradiction. Hence \(l\) is a measure on \(\mathcal{I}\). The result now follows from Carathéodory.

Surprisingly, the extension of this result to uncountable products is easy. Indeed, for any set \(A\) in the \(\sigma\)-algebra generated by cylinder sets, there is a countable \(I\) such that \(A\) is also in the \(\sigma\)-algebra generated by cylinder sets \(\prod A_i\) with \(A_i = X_i\) for \(i \notin I\). Thus the measure need only be defined on countable products.
Definition A function \( f: (X, \mathcal{A}) \to (Y, \mathcal{B}) \) between measurable spaces is called measurable if for all \( B \in \mathcal{B} \), \( f^{-1}[B] \in \mathcal{A} \).

Definition A function \( f: (X, \mathcal{A}) \to \mathbb{R}^* \) is measurable iff it is measurable with respect to the Borel \( \sigma \)-algebra on \( \mathbb{R}^* \).

Note: we do not in general use complete measures on \( Y \) since this may make many ‘nice’ functions non-measurable. In particular, if we use Lebesgue measurable sets then there exist continuous functions that are not measurable: Take two Cantor-like sets with \( \lambda(C_1) > 0 = \lambda(C_2) \) and construct a continuous bijection \( f: [0, 1] \to [0, 1] \), \( f[C_1] = C_2 \), by making it map each interval of \([0, 1] \setminus C_1 \) linearly onto the corresponding interval of \([0, 1] \setminus C_2 \). Then any non-measurable subset \( E \subseteq C_1 \) is the inverse image of the measurable set \( f(E) \subseteq C_2 \).

Since \( \{ B : f^{-1}[B] \in \mathcal{A} \} \) is a \( \sigma \)-algebra on \( Y \), it is enough to check the condition on any set of \( B \)’s that generate \( \mathcal{B} \) as a \( \sigma \)-algebra. In particular, \( f: X \to \mathbb{R}^* \) is measurable iff \( f^{-1}((a, \infty]] \) is measurable for all \( a \in \mathbb{R} \), or even just all \( a \in \mathbb{Q} \).

Lemma 1. For functions \( (X, \mathcal{A}) \to \mathbb{R}^* \)

1. If \( (X, \mathcal{A}) = (\mathbb{R}, \mathcal{B}) \) or \( (\mathbb{R}, \mathcal{L}) \), then any continuous function is measurable.
2. The characteristic function \( \chi_S \) is measurable iff \( S \in \mathcal{A} \).
3. If \( f_n \) are measurable then \( \sup_n f_n, \inf_n f_n, \liminf f_n, \) and \( \limsup f_n \) are measurable.
4. If \( f, g \) are measurable then \( f + g, f - g, fg \) and \( f/g \) are measurable as functions on the set where they are defined. The set where they are defined is also measurable.

Definition A simple function is a measurable function \( \phi: X \to \mathbb{R} \) such that \( \phi[X] \) is finite. Equivalently \( \phi = \sum_{i=1}^n a_i \chi_{S_i} \) where \( S_i \) are measurable subsets of \( X \), \( a_i \in \mathbb{R} \), and \( \chi_S \) is the characteristic function of \( S \). We may choose the \( S_i \) to be disjoint.

Lemma 2. If \( f: X \to [0, \infty] \) is measurable, then there exists an increasing sequence of simple functions \( 0 \leq \phi_1 \leq \phi_2 \leq \ldots \) with \( \phi_n \to f \) pointwise.

Lemma 3. If \( f: X \to Y \) is any function and \( (Y, \mathcal{B}) \) is a measurable space, then \( \sigma(f) = \{ f^{-1}[B] : B \in \mathcal{B} \} \) is a \( \sigma \)-algebra on \( X \).

We call \( \sigma(f) \) the \( \sigma \)-algebra on \( X \) generated by \( f \). The function \( f: (X, \mathcal{A}) \to (Y, \mathcal{B}) \) is measurable iff \( \sigma(f) \subseteq \mathcal{A} \).

More generally, if \( f_1, f_2, \ldots \) are functions on \( X \) to a measurable space, \( \sigma(f_1, f_2, \ldots) \) is the \( \sigma \)-algebra generated by all the \( \sigma(f_i) \)’s and is the smallest \( \sigma \)-algebra on \( X \) making all the \( f_i \) measurable.

Example The (uncompleted) \( \sigma \)-algebra defined on a product space (finite or infinite) is just \( \sigma(\pi_1, \pi_2, \ldots) \) where \( \pi_i \) is the projection map onto the \( i \)'th coordinate.
Math 7351  8. Integration  Spring 2005

**Lemma** If \( f : X \to [0, \infty] \) is a measurable function, then the shadow of \( f \), \( S(f) = \{(x, y) : 0 \leq y < f(x)\} \) is a \((\mu \times \lambda)\)-measurable subset of \( X \times \mathbb{R} \).

**Proof.** Clear for simple functions, and \( S(f) = \bigcup S(\phi_n) \) where \( \phi_1 \leq \phi_2 \leq \ldots, \phi_n \to f \).

**Definition** If \( f : X \to [0, \infty] \) is measurable, the integral of \( f \) is \( \int f d\mu = (\mu \times \lambda)(S(f)) \).

**Theorem (Dominated Convergence Theorem)** If \( f : X \to \mathbb{R}^* \) is measurable and \( \int |f| d\mu < \infty \) then we say \( f \) is integrable and define \( \int f d\mu = \int f^+ d\mu - \int f^- d\mu \) where \( f^+(x) = \max\{f(x), 0\} \), \( f^-(x) = \max\{-f(x), 0\} \).

Clearly \( \int \phi d\mu = \sum_{i=1}^{n} a_i \mu(S_i) \) for any simple non-negative \( \phi = \sum_{i=1}^{n} a_i \chi_{S_i} \).

**Theorem (Monotone Convergence Theorem)** If \( 0 \leq f_1 \leq f_2 \leq \ldots \) is an increasing sequence of non-negative measurable functions on \( X \), then \( \int \lim f_n d\mu = \lim \int f_n d\mu \).

**Proof.** \( S(f_1) \subseteq S(f_2) \subseteq \ldots \) and \( S(f) = \bigcup S(f_n) \).

**Corollary** If \( f : X \to [0, \infty] \) is measurable then \( \int f d\mu = \sup_{\phi} \int \phi d\mu \) where the supremum is taken over simple \( \phi \) with \( 0 \leq \phi \leq f \).

**Proof.** \( S(\phi) \subseteq S(f) \), so \( \int \phi \leq \int f \), and if \( 0 \leq \phi_1 \leq \ldots, \phi_n \to f \), then \( \int \phi_n \to \int f \).

Note, this gives an alternative definition of the integral, and shows that it does not depend on the choice of \( \mu \times \lambda \) when \( \mu \) is not \( \sigma \)-finite.

**Theorem** Suppose \( f, g : X \to [0, \infty] \) are measurable, (resp. \( f, g : X \to \mathbb{R}^* \) integrable).

1. If \( f \leq g \) then \( \int f \, d\mu \leq \int g \, d\mu \)
2. If \( c \geq 0 \) (resp. \( c \in \mathbb{R} \)) then \( \int cf \, d\mu = c \int f \, d\mu \)
3. \( \int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu \)
4. If \( f \geq 0 \) then \( \int f \, d\mu = 0 \) iff \( f = 0 \) a.e.

**Proof.** For 2 and 3 with \( f, g \geq 0 \) prove it first with simple functions and take limits. For 4, \( \Rightarrow, \int f \, d\mu \leq \frac{1}{N} \mu\{x : f(x) > \frac{1}{N}\} \) and \( \{x : f(x) > 0\} = \bigcup\{x : f(x) > \frac{1}{N}\} \).

**Theorem (Fatou’s Lemma)** If \( f_i \geq 0 \) are non-negative measurable functions then \( \int \lim f_n \, d\mu \leq \lim \int f_n \, d\mu \).

**Proof.** If \( g_n = \inf_{r \geq n} f_r \), then \( g_n \) is increasing and \( \int \lim f_n \, d\mu = \lim \int g_n \, d\mu = \lim \int g_n \, d\mu \leq \lim_{n} \inf_{r \geq n} \int f_r \, d\mu = \lim \int f_n \, d\mu \).

**Theorem (Dominated Convergence Theorem)** If \( g \) is integrable and \( |f_n| \leq g \) and \( f_n \) converges pointwise then \( \int \lim f_n = \lim \int f_n \).

**Proof.** Apply Fatou to \( g - f_n \) and \( g + f_n \).

Definition A collection $\mathcal{M}$ of subsets of $X$ is a monotone class if whenever $A_1 \subseteq A_2 \subseteq \ldots$, $A_i \in \mathcal{M}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{M}$ and whenever $A_1 \supseteq A_2 \supseteq \ldots$, $A_i \in \mathcal{M}$, then $\bigcap_{i=1}^{\infty} A_i \in \mathcal{M}$.

Lemma 1. If $\mathcal{A}$ is an algebra, then the smallest monotone class $\mathcal{M}$ containing $\mathcal{A}$ is equal to the $\sigma$-algebra $\sigma(\mathcal{A})$ generated by $\mathcal{A}$.

Proof. The intersection of all monotone classes $\supseteq \mathcal{A}$ is a monotone class, so $\mathcal{M}$ exists. Let $M(\mathcal{A}) = \{B \subseteq X : A \cup B, A \setminus B, B \setminus A \in \mathcal{M}\}$. Then $M(\mathcal{A})$ is a monotone class. If $A \in \mathcal{A}$ then $A \subseteq M(\mathcal{A})$, so $\mathcal{M} \subseteq M(\mathcal{A})$. But then (reversing the roles of $A$ and $B$), if $A \in M(\mathcal{A})$ then $A \subseteq M(\mathcal{A})$, so $\mathcal{M} \subseteq M(\mathcal{A})$. But then $\mathcal{M}$ is closed under finite unions and differences, so is a ring. If $A_i \in \mathcal{M}$, then $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} (\bigcup_{i=1}^{\infty} A_i) \in \mathcal{M}$ and as $X \in \mathcal{A} \subseteq M$, $\mathcal{M}$ is a $\sigma$-algebra. Thus $\sigma(\mathcal{A}) \subseteq \mathcal{M}$, but $\sigma(\mathcal{A})$ is a monotone class, so $\sigma(\mathcal{A}) = \mathcal{M}$.

Lemma 2. Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are $\sigma$-finite measure spaces. If $E$ is a $\mu \times \nu$-measurable set and $g(x) = \nu(E_x)$, then $g$ exists, is $\mu$-measurable, and $\mu \times \nu(E) = \int g \, d\mu$.

Proof. First assume $\mu$ and $\nu$ are finite. Consider $\mathcal{M} = \{E \subseteq X \times Y : g(x) = \nu(E_x) \text{ exists}, \text{is measurable}, \text{and } \mu \times \nu(E) = \int g \, d\mu\}$. Then $\mathcal{M}$ contains all measurable rectangles, and is closed under finite disjoint unions, so contains the algebra generated by measurable rectangles. But $\mathcal{M}$ is a monotone class (use the fact that $\mu$ and $\nu$ are finite, and the DCT for $\int g \, d\mu$). Thus $\mathcal{M} \subseteq \mathcal{A} \otimes \mathcal{B}$. For $\sigma$-finite $\mu$ and $\nu$, write $X \times Y$ as a union of increasing finite rectangles $X_i \times Y_i$, prove the result for $E \cap (X_i \times Y_i)$ and take limits.

Corollary 3. If $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are complete $\sigma$-finite measure spaces, $E$ is $\mu \times \nu$-measurable, and $g(x) = \nu(E_x)$, then $g$ exists $\mu$-a.e., is $\mu$-measurable, and $\mu \times \nu(E) = \int g \, d\mu$.

Proof. If $E$ is $\mu \times \nu$-measurable and $\mu \times \nu(E) = 0$ then by Lemma 2, $\int g \, d\mu = 0$, so $g = 0$ a.e.. Thus if $E$ is a subset of a set of $\mu \times \nu$-measure zero then $g = 0$ a.e., and the result holds. Writing $E$ as a union of a $\mu \times \nu$-measurable set and a subset of a set with $\mu \times \nu$-measure zero gives the result.

Theorem (Fubini-Tonelli) If $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are $\sigma$-finite measure spaces and $f : X \times Y \to \mathbb{R}^+$ is non-negative and $\mu \times \nu$-measurable, then $f(x, :)$ : $Y \to \mathbb{R}^+$ is $\nu$-measurable for all $x \in X$, $g(x) = \int f(x, y) \, d\nu$ is $\mu$-measurable and $\int f(x, y) \, d\mu \times \nu = \int \int f(x, y) \, d\mu \, d\nu$. Similarly, if $f$ is $\mu \times \nu$-integrable, then $f(x, :)$ is $\nu$-integrable for $\mu$-a.e. $x \in X$, $g(x) = \int f(x, y) \, d\nu$ is $\mu$-integrable and $\int f(x, y) \, d\mu \times \nu = \int \int f(x, y) \, d\mu \, d\nu$.

Proof. Lemma 2 shows that this holds when $f = \chi_E$. Thus by linearity it holds for simple functions, and MCT implies it holds for all non-negative measurable $f$. For integrable $f$, apply result to $f^+, |f|$, and use $\int \int |f| \, d\mu \, d\nu < \infty$ to show $f(x, :)$ is $\nu$-integrable $\mu$-a.e.. A corresponding result holds for $\mu \times \nu$ provided $\mu$ and $\nu$ are complete and we replace ‘all $x \in X$’ with ‘$\mu$-a.e. $x \in X$’.

Example $\int_0^\infty \int_0^\infty \frac{x^2-y^2}{(x^2+y^2+1)^2} \, dx \, dy \neq \int_0^\infty \int_0^\infty \frac{x^2-y^2}{(x^2+y^2+1)^2} \, dy \, dx$. 
Math 7351 10. Radon-Nikodym Spring 2005

**Definition** If \( \mu \) and \( \nu \) are two signed measures on a measurable space \((X, A)\) then we say \( \nu \) is absolutely continuous with respect to \( \mu \), \( \nu \ll \mu \), iff every \( \mu \)-null set is \( \nu \)-null. We say \( \nu \) is singular with respect to \( \mu \), \( \nu \perp \mu \), if they are mutually singular, i.e., \( \nu \) is supported on a \( \mu \)-null set.

**Theorem (Radon-Nikodym)** Let \( \mu \) and \( \nu \) be (positive) measures on the same measurable space \((X, A)\), with \( \nu \ll \mu \) and \( \mu \) \( \sigma \)-finite. Then there exists a measurable function \( f : X \to [0, \infty] \) such that \( \nu(A) = \int_A f \, d\mu \) for all \( A \in A \). Moreover if \( f \) and \( g \) are two such functions then \( f = g \) \( \mu \)-a.e..

**Proof.** Assume first that \( \mu \) is finite. Then for all \( \alpha \in \mathbb{Q} \), \( \alpha \geq 0 \), \( \nu - \alpha \mu \) is a signed measure. Let \( X = X_0^+ \cup X_0^- \) be a corresponding Hahn decomposition. We may assume \( X_0^+ = X \).

Fix a measurable \( E \), and \( N > 0 \), and let \( E_i = E \cap f^{-1}[\{\frac{i}{N}, \frac{i+1}{N}\}] \). Then \( E_i \subseteq X_i^- \), so \( \nu(E_i) \leq \frac{i+1}{N} \mu(E_i) \). Also \( E_i \subseteq X_i^+ \cup \bigcup_{\alpha > \beta} (X_\alpha^+ \setminus X_\beta^+) \) for all \( \beta < \frac{i}{N} \). Thus \( \nu(E_i) \geq \nu(E_i \cap X_\beta^+) \geq \beta \mu(E_i \cap X_\beta^+) = \beta \mu(E_i) \). Thus \( \nu(E_i) \geq \frac{\beta}{N} \mu(E_i) \). But \( \frac{\beta}{N} \leq f \leq \frac{\beta+1}{N} \) on \( E_i \), so \( \frac{\beta}{N} \mu(E_i) \leq \int_{E_i} f \, d\mu \leq \frac{\beta+1}{N} \mu(E_i) \). Thus

\[-\frac{1}{N} \mu(E_i) \leq \nu(E_i) - \int_{E_i} f \, d\mu < \frac{1}{N} \mu(E_i) \]  

If we let \( E_\infty = E \cap f^{-1}[\{\infty\}] \) then \( E \setminus E_\infty = \bigcup E_i \) is a disjoint union. Thus by MCT and countable additivity of \( \nu \) and \( \mu \),

\[-\frac{1}{N} \mu(E \setminus E_\infty) \leq \nu(E \setminus E_\infty) \leq \int_{E \setminus E_\infty} f \, d\mu \leq \frac{1}{N} \mu(E \setminus E_\infty) \]  

Since this holds for all \( N \) and \( \mu(E) < \infty \), \( \nu(E \setminus E_\infty) = \int_{E \setminus E_\infty} f \, d\mu = \infty \). Finally, if \( \mu(E_\infty) > 0 \) then \( \nu(E_\infty) > \alpha \mu(E_\infty) \) for arbitrarily large \( \alpha \)'s, so \( \nu(E_\infty) = \int_{E_\infty} f \, d\mu = \infty \). On the other hand, if \( \mu(E_\infty) = 0 \) then \( \nu(E_\infty) = 0 \) since \( \nu \ll \mu \), and \( \nu(E_\infty) = \int_{E_\infty} f \, d\mu = 0 \). Thus by addition \( \nu(E) = \int_E f \, d\mu \).

For \( \sigma \)-finite \( \mu \), write \( X = \bigcup X_i \) with \( \mu(X_i) < \infty \) and disjoint. We can define \( f_i \) on \( X_i \) by \( \nu(A \cap X_i) = \int_{A \cap X_i} f_i \, d\mu \). Now let \( f = \sum f_i \chi_{X_i} \) and use MCT. For uniqueness, let \( E = \{ x : f(x) - g(x) > \frac{1}{n} \text{ and } g(x) < n \} \cap X_i \). Then \( \nu(E) = \int f \, d\mu \geq \frac{1}{n} \mu(E) + \int g \, d\mu = \nu(E) \), which implies \( \mu(E) = 0 \) (note that \( f g \, d\mu < \infty \)). Taking unions over all \( n \) and \( i \) we get \( \mu(\{ x : f(x) > g(x) \}) = 0 \) and similarly \( \mu(\{ x : f(x) < g(x) \}) = 0 \). Thus \( f = g \) \( \mu \)-a.e.. \( \Box \)

**Definition** We define a Radon-Nikodym derivative of \( \nu \) with respect to \( \mu \), \( \frac{d\nu}{d\mu} \), to be this \( f \). Note that it is only defined up to equality \( \mu \)-a.e..

Note that if \( f \) is any non-negative measurable function then \( \nu(E) = \int_E f \, d\mu \) defines a measure with \( \nu \ll \mu \) and Radon-Nikodym derivative \( \frac{d\nu}{d\mu} = f \) \( \mu \)-a.e..
Corollary (Lebesgue Decomposition) \If \((X, A, \mu)\) is a \(\sigma\)-finite measure space, then any \(\sigma\)-finite measure \(\nu\) on \((X, A)\) can be written in the form \(\nu = \nu_c + \nu_s\) where \(\nu_c \ll \mu\) and \(\nu_s \perp \mu\).

Proof. Let \(\psi = \nu + \mu\), then \(\psi\) is \(\sigma\)-finite and \(\mu \ll \psi\). Write \(\mu(E) = \int_E f \, d\psi\) and let \(X = A \cup B\) where \(A = \{x : f(x) > 0\}\) and \(B = \{x : f(x) = 0\}\). Define \(\nu_c(E) = \nu(E \cap A)\) and \(\nu_s(E) = \nu(E \cap B)\). Then \(\nu = \nu_c + \nu_s\), \(\nu_s\) is supported on \(B\) and \(\mu(B) = 0\), so \(\nu_s \perp \mu\). If \(\mu(E) = 0\) then \(\psi(E \cap A) = 0\), so \(\nu_c(E) = \nu(E \cap A) \leq \psi(E \cap A) = 0\), and \(\nu_c \ll \mu\). \(\square\)

Recall the Lebesgue-Stieltjes measure on \((\mathbb{R}, B)\) given by \(\mu_F((a, b]) = F(b) - F(a)\) for some increasing right-continuous function \(F\). We generally denote the integral with respect to \(\mu_F\) by \(\int f(x) \, dF\).

Theorem The Lebesgue-Stieltjes measure \(\mu_F\) is absolutely continuous with respect to Lebesgue measure \(\lambda\) iff \(F\) is an absolutely continuous function. In this case \(\frac{d\mu_F}{d\lambda} = F'\) \(\lambda\)-a.e..

Proof. If \(\mu_F \ll \lambda\) then by Radon-Nikodym, \(F(b) - F(a) = \mu_F((a, b]) = \int_{(a, b]} \frac{d\mu_F}{d\lambda} \, d\lambda\). But then \(F(x) = F(a) + \int_a^x \frac{d\mu_F}{d\lambda}(t) \, dt\) and \(\frac{d\mu_F}{d\lambda}(t) \geq 0\) is measurable, so \(F(x)\) is absolutely continuous. Conversely, suppose \(F\) is absolutely continuous, then \(F'\) exists a.e., and \(F(b) - F(a) = \int_a^b F'(x) \, dx\). Define \(\mu(E) = \int_E F'(x) \, dx\). Then \(\mu\) is a measure on \((\mathbb{R}, B)\) and \(\mu((a, b]) = \mu_F((a, b])\). Let \(\mathcal{M} = \{E \subseteq (n, n+1] : \mu(E) = \mu_F(E)\}\). Then \(\mathcal{M}\) is a monotone class (using \(\mu_F((n, n+1]) < \infty\) for decreasing limits). But \(\mathcal{M}\) also contains \((a, b]\) for \(n \leq a < b \leq n + 1\) and is closed under finite disjoint unions, so contains the algebra on \((n, n + 1]\) generated by half-open intervals. Thus \(\mathcal{M}\) contains all Borel sets in \((n, n + 1]\). Finally, for arbitrary \(E \in B\), \(\mu_F(E) = \sum_n \mu_F(E \cap (n, n+1]) = \sum_n \mu(E \cap (n, n+1]) = \mu(E)\), so \(\mu = \mu_F\). But then \(\mu_F(E) = \int_E F'(x) \, dx\), so if \(\lambda(E) = 0\) then \(\mu_F(E) = 0\), so \(\mu_F \ll \lambda\). Finally, \(F' = \frac{d\mu_F}{d\lambda}\) \(\lambda\)-a.e. by uniqueness of the Radon-Nikodym derivative. \(\square\)

Exercises

1. If \(\nu \ll \mu\), then \(\int f \, d\nu = \int f \frac{d\nu}{d\mu} \, d\mu\).

2. If \(\psi \ll \nu \ll \mu\), then \(\frac{d\psi}{d\mu} = \frac{d\psi}{d\nu} \frac{d\nu}{d\mu} \mu\)-a.e.

3. If \(\psi, \nu \ll \mu\), then \(\frac{d(\psi + \nu)}{d\mu} = \frac{d\psi}{d\mu} + \frac{d\nu}{d\mu} \mu\)-a.e.

4. Extend the Radon-Nikodym theorem to the case when \(\nu\) is a signed measure.
One can construct a model of probability using measure theory. The measure space $(X, \mathcal{A}, \mu)$ is usually denoted $(\Omega, \mathcal{A}, \mathbb{P})$, where $\Omega$ is the sample space, or the set of possible outcomes. The $\sigma$-algebra $\mathcal{A}$ is the set of all events, and $\mathbb{P}$ is a probability measure which assigns to each event $E \in \mathcal{A}$ a probability $\mathbb{P}(E) \in [0,1]$. An event occurs almost surely or a.s., if $\mathbb{P}(E) = 1$, or equivalently $\mathbb{P}(\text{not } E) = 0$.

A random variable is a measurable function on $\Omega$ (usually to $\mathbb{R}$ and usually denoted in upper case $X, Y, \ldots$, lower case variables typically denote constants). We write, for example, $\mathbb{P}(X > c)$ as a shorthand for $\mathbb{P}\{\omega \in \Omega : X(\omega) > c\}$. The $\sigma$-algebra $\sigma(X) = \{X^{-1}[B] : B \text{ Borel}\}$ is the set of events that can be described in terms of the value of $X$ as $'X \in B'$. If one is given random variables $\sigma$ is independent of the other variables. Two random variables $X, Y$ are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$ for any Borel set $B$. The cumulative distribution function of a random variable is the function $F(x) = \mathbb{P}(X \leq x)$. The measure $\mu$ is just the Lebesgue-Stieltjes measure corresponding to $F$. If $F$ is absolutely continuous, then $f = F'$ is called the probability density function of $X$, and is just the Radon-Nikodym derivative $\frac{d\mu}{dx}$. Note that $\mathbb{E} = \int x dF = \int xf(x) dx$ when defined.

Any real-valued random variable gives rise to a probability measure on $(\mathbb{R}, \mathcal{B})$ by setting $\mu(B) = \mathbb{P}(X \in B)$ for any Borel set $B$. The cumulative distribution function of a random variable is the function $F(c) = \mathbb{P}(X \leq c)$. The measure $\mu$ is just the Lebesgue-Stieltjes measure corresponding to $F$. If $F$ is absolutely continuous, then $f = F'$ is called the probability density function of $X$, and is just the Radon-Nikodym derivative $\frac{d\mu}{dx}$. Note that $\mathbb{E} = \int x dF = \int xf(x) dx$ when defined.

If $\mathcal{A}_1$ and $\mathcal{A}_2$ are two sub-$\sigma$-algebras of $\mathcal{A}$, we say $\mathcal{A}_1$ and $\mathcal{A}_2$ are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$ for all $A \in \mathcal{A}_1, B \in \mathcal{A}_2$. Two events $A$ and $B$ are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$, or equivalently the $\sigma$-algebras generated by $\{A\}$ and $\{B\}$ are independent. Two random variables $X$ and $Y$ are independent if $\sigma(X)$ and $\sigma(Y)$ are independent. In other words, any event describable in terms of $X$ is independent of any event describable in terms of $Y$. More generally, any number of $\sigma$-algebras $\mathcal{A}_i$ are independent if each $\mathcal{A}_i$ is independent of the $\sigma$-algebra generated by all the others, and similarly for events and random variables. If one is given random variables $X_i$ on different probability spaces $(\Omega_i, \mathcal{A}_i, \mathbb{P}_i)$, one can construct a probability space on which all the $X_i$ are independent by taking the product space with the product measure.

**Warning:** Suppose $X_1, \ldots X_{n-1}$ are independent random variables that take the values 0 or 1 each with probability $\frac{1}{2}$. Let $X_n \in \{0, 1\}$ be the sum $X_1 + \cdots + X_{n-1} \mod 2$. Then any subset of the $X_i$'s of size $< n$ are independent, but $X_1, \ldots, X_n$ are not independent.

**Exercises**

1. If $X_1, X_2, \ldots$ are random variables with $\sum \mathbb{E}|X_i| < \infty$ then $\mathbb{E}(\sum X_i) = \sum \mathbb{E}(X_i)$.

2. If $X_1, \ldots, X_n$ are independent random variables with $\mathbb{E}|X_i| < \infty$ then $\mathbb{E}(\prod_{i=1}^n X_i) = \prod_{i=1}^n \mathbb{E}(X_i)$ and $\text{Var}(\sum X_i) = \sum \text{Var}(X_i)$.

3. Tchebychev’s Inequality: If $\mathbb{E}|X| < \infty$ and $t > 0$ then $\mathbb{P}(|X - \mathbb{E}X| \geq t) \leq \text{Var}(X)/t^2$. Spring 2005
Theorem (Kolmogorov’s 0–1 law) Suppose $X_1, X_2, \ldots$ are independent random variables and $E$ is a tail event, i.e., an event such that for all $n$, $E$ only depends on the values of $X_{n+1}, X_{n+2}, \ldots$. Then $\mathbb{P}(E) = 0$ or 1.

Proof. The set $\mathcal{M}$ of all events that are independent of $E$ is a monotone class: If $\mathbb{P}(E \cap A_i) = \mathbb{P}(E) \mathbb{P}(A_i)$ and $A_i \in \mathcal{M}$ is a monotonic sequence, then the limit $A = \bigcup A_i$ or $\bigcap A_i$ satisfies $\mathbb{P}(E \cap A) = \lim \mathbb{P}(E \cap A_i) = \mathbb{P}(E) \lim \mathbb{P}(A_i) = \mathbb{P}(E) \mathbb{P}(A)$, so $A \in \mathcal{M}$. Now $E \in \sigma(X_{n+1}, X_{n+2}, \ldots)$, so $E$ is independent of $\sigma(X_1, \ldots, X_n)$. Thus $\mathcal{C} = \bigcup_n \sigma(X_1, \ldots, X_n) \subseteq \mathcal{M}$. But $\mathcal{C}$ is an algebra (check this), so $\mathcal{M}$ contains the $\sigma$-algebra generated by $\mathcal{C}$, which is just $\sigma(X_1, X_2, \ldots)$. But then $E \in \mathcal{M}$, so $E$ is independent of $E$. But then $\mathbb{P}(E) = \mathbb{P}(E \cap E) = \mathbb{P}(E) \mathbb{P}(E)$, so $\mathbb{P}(E) = 0$ or 1. \qed

Example Events such as ‘$\lim X_i \leq c$’ and ‘$\lim \frac{1}{n} \sum_{i=1}^{n} X_i = c$’ are tail events.

Example Consider $\mathbb{Z}^2$ and join neighboring (horizontally or vertically adjacent) points independently with probability $p$. Then the probability that there is an infinite connected subset of $\mathbb{Z}^2$ is either 0 or 1. (In fact it is 1 for $p > 0.5$ and 0 for $p \leq 0.5$, but this is very much harder to prove).

Conditional Expectation

In elementary probability theory, one defines the conditional probability of $A$ given $B$ as $\mathbb{P}(A \mid B) = \mathbb{P}(A \cap B) / \mathbb{P}(B)$. This works as long as $\mathbb{P}(B) > 0$. But there are many instances when we would like to apply conditional probability when $\mathbb{P}(B) = 0$. More specifically, if $Z$ is a random variable, we would like to define $\mathbb{P}(A \mid Z = z)$ as a function $\phi(z)$ even when $\mathbb{P}(Z = z) = 0$. If we consider $\mathbb{P}(A \mid Z) = \phi(Z)$, then what we are asking for is a new random variable that depends only on the value of $Z$, i.e., is $\sigma(Z)$-measurable.

We first define conditional expectation. Given a $\sigma$-algebra $\mathcal{A}_0 \subseteq \mathcal{A}$ and an integrable random variable $X$ ($\mathbb{E}[X] < \infty$), define for $A \in \mathcal{A}_0$, $\mu(A) = \mathbb{E}(I_A X)$. Then $\mu$ is a signed measure on $(\Omega, \mathcal{A}_0)$. Also, $\mu \ll \mathbb{P}$, so by the Radon-Nikodym theorem, there exists an $\mathcal{A}_0$-measurable $Y$ such that $\mathbb{E}(I_A X) = \mathbb{E}(I_A Y)$ for all $A \in \mathcal{A}_0$. This random variable $Y$ is denoted $\mathbb{E}(X \mid \mathcal{A}_0)$ and is called the conditional expectation of $X$ given $\mathcal{A}_0$. It is only defined up to equality a.s.. We define, for example, $\mathbb{E}(X \mid Y, Z)$ to be $\mathbb{E}(X \mid \sigma(Y, Z))$. Conditional probability is defined by, for example, $\mathbb{P}(E \mid \mathcal{A}_0) = \mathbb{E}(1_E \mid \mathcal{A}_0)$.

Lemma Assuming all relevant quantities are defined,

1. $\mathbb{E}(X \mid Y) = \phi(Y)$ a.s. for some Borel measurable $\phi : \mathbb{R} \rightarrow \mathbb{R}$,
2. if $X$ and $\mathcal{A}_0$ are independent then $\mathbb{E}(X \mid \mathcal{A}_0) = \mathbb{E}X$ a.s.,
3. if $X$ is $\mathcal{A}_0$-measurable then $\mathbb{E}(XY \mid \mathcal{A}_0) = X \mathbb{E}(Y \mid \mathcal{A}_0)$ a.s.,
4. if $\mathcal{A}_1 \subseteq \mathcal{A}_0 \subseteq \mathcal{A}$ then $\mathbb{E}(X \mid \mathcal{A}_1) = \mathbb{E}(\mathbb{E}(X \mid \mathcal{A}_0) \mid \mathcal{A}_1)$ a.s., in particular $\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X \mid \mathcal{A}_0))$. 
Math 7351 12. $L_p$ spaces Spring 2005

Suppose $(X, \mathcal{A}, \mu)$ is a measure space and $f: X \to \mathbb{R}^*$ is measurable. Define $\|f\|_p = \left(\int |f|^p \, d\mu\right)^{1/p}$ for $1 \leq p < \infty$ and $\|f\|_\infty = \text{ess sup} \{f\} = \inf\{c: \mu\{x: |f(x)| > c\} = 0\}$.

**Lemma** $f \equiv g$ a.e. $\Rightarrow \|f\|_p = \|g\|_p$ and $\|f\|_p = 0$ iff $f \equiv 0$ a.e.

**Theorem (Minkowski)** $\|f + g\|_p \leq \|f\|_p + \|g\|_p$

Proof. $|x|^p$ convex $\Rightarrow \left(\frac{\|f\|_p}{\|f\| + \|g\|} + \frac{\|g\|}{\|f\| + \|g\|}\right)^p \leq \left(\|f\|_p + \|g\|_p\right)^p$. Now $f$. □

Define $L^p(X, \mathcal{A}, \mu)$ to be $\{f: \|f\|_p < \infty\}/\sim$, where $f \sim g$ iff $f = g$ a.e..

**Lemma** $L^p(X, \mathcal{A}, \mu)$ is a vector space, and $\|\cdot\|_p$ induces a norm on $L^p(X, \mathcal{A}, \mu)$.

**Theorem (Riesz-Fischer)** $L^p(X, \mathcal{A}, \mu)$ is complete wrt $\|\cdot\|_p$, so is a Banach space.

Proof. First show that $L^p$ is complete iff $\sum \|f_n\|_p < \infty \Rightarrow \sum f_n$ converges in $L^p$.

Now $\sum \|f_n\|_p < \infty$ gives $g(x) = \sum |f_n(x)| \in L^p$ by MCT, so $g < \infty$ a.e., and $f(x) = \sum f_n(x)$ converges a.e.. Apply DCT to show $\|f - \sum f_n\|_p \to 0$, (dominate with $|g|^p$). □

**Theorem (Hölder)** If $\frac{1}{p} + \frac{1}{q} = 1$, $f \in L^p$, $g \in L^q$ then $\int |fg| \, d\mu \leq \|f\|_p \|g\|_q$.

Proof. Use Young’s inequality $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ with $a = \frac{|f|}{\|f\|_p}$, $b = \frac{|g|}{\|g\|_q}$. Now $f$. □

**Lemma** For $p < \infty$, simple $L^p$ functions are dense in $L^p$.

Proof. Let $0 \leq \phi_1 \leq \phi_2 \leq \cdots \to |f|$, then $\psi_n = \phi_n \text{ sgn} f$ is simple, $\|\psi_n\|_p = \|\phi_n\|_p \leq \|f\|_p$ and $\|f - \psi_n\|_p \to 0$ by DCT (dominate by $|f|^p$). □

**Lemma** For $p < \infty$, the support $\text{supp} f = \{x: f(x) \neq 0\}$ of any $f \in L^p$ is $\sigma$-finite.

Proof. $\text{supp} f = \bigcup_n \{x: |f(x)| > \frac{1}{n}\}$, and $\mu\{x: |f(x)| > \frac{1}{n}\} \leq \int (n|f|^p) = n^p\|f\|_p^p < \infty$. □

**Theorem (Riesz Representation Theorem)** Let $F$ be a bounded linear functional on $L^p(X, \mathcal{A}, \mu)$, $1 \leq p < \infty$ and suppose either $p > 1$ or $(X, \mathcal{A}, \mu) \sigma$-finite. Then there is a unique function $g \in L^q(X, \mathcal{A}, \mu)$, $\frac{1}{p} + \frac{1}{q} = 1$, such that

$$F(f) = \int f g \quad \text{for all } f \in L^p(X, \mathcal{A}, \mu).$$

Moreover, for all such $g$, the above formula defines a linear functional with $\|F\| = \|g\|_q$.

Proof. Assume first that $\mu$ is finite. Now $\chi_E \in L^p$ for any $E$. Define $\nu(E) = F(\chi_E) \in \mathbb{R}$.

Claim 1: $\nu$ is a finite signed measure.

Finite clear. If $E = \bigcup E_i$ is disjoint, then $\forall N \geq n_0: \mu(E \setminus F_N) < \epsilon$ where $F_N = \bigcup_{i=1}^N E_i$. 

Therefore, $\nu(E) = \sum \nu(E_i)$, where $\epsilon \mu(E_i)$ converges a.e. for any $\epsilon > 0$. □

Moreover, for all such $g$, the above formula defines a linear functional with $\|F\| = \|g\|_q$. □

Proof. Assume first that $\mu$ is finite. Now $\chi_E \in L^p$ for any $E$. Define $\nu(E) = F(\chi_E) \in \mathbb{R}$.

Claim 1: $\nu$ is a finite signed measure.

Finite clear. If $E = \bigcup E_i$ is disjoint, then $\forall N \geq n_0: \mu(E \setminus F_N) < \epsilon$ where $F_N = \bigcup_{i=1}^N E_i$, 

Moreover, for all such $g$, the above formula defines a linear functional with $\|F\| = \|g\|_q$. □
so \(|\nu(E) - \sum_{i=1}^{N} \nu(E_i)| = |F(\chi_E - \sum_{i=1}^{N} \chi_{E_i})| = |F(\chi_{E \setminus F_N})| \leq \|F\| \|\chi_{E \setminus F_N}\|_p \leq \|F\| \epsilon^{1/p}\). Hence \(\nu(E) = \sum_{i=1}^{\infty} \nu(E_i)\).

Claim 2: \(\nu \ll \mu\).

\(\mu(E) = 0 \Rightarrow |\nu(A)| = |F(\chi_A)| \leq \|F\| \|\chi_A\|_p = 0\) for any \(A \subseteq E\).

Now by the Radon-Nikodym theorem \(F(\chi_E) = \nu(E) = \int_E g \, d\mu = \int \chi_E g \, d\mu\). So by linearity, \(F(\phi) = \int \phi g \, d\mu\) for any simple function \(\phi\).

Claim 3: \(\|g\|_q \leq \|F\|\), in particular \(g \in L^q\).

Let \(0 \leq \phi_1 \leq \phi_2 \leq \cdots \to |g|^{p/q}\). Then \(\|\phi_n\|_p = \int \phi_n^p = \int \phi_n^{p/q+1} \leq \int |g| \phi_n = F(\phi_n \text{sgn } g) \leq \|F\| \|\phi_n\|_p\). Hence \(\|\phi_n\|_p \leq \|F\|\), and so \(\int \phi_n^p \to \int |g|^q\) by MCT, so \(\|g\|_q \leq \|F\|\). For \(q = \infty\) let \(E = \{x: |g(x)| > c\}\), then \(c\mu(E) \leq \int_E |g| = F(\chi_E \text{sgn } g) \leq \|F\| \|\chi_E\|_1 = \|F\| \mu(E)\), so if \(c > \|F\|\) then \(\mu(E) = 0\).

Claim 4: \(F(f) = \int f g\).

Let \(\phi_n \to f\) in \(L^p\), then \(\|F(f) - \int f g\| \leq \|F(f) - \phi_n g\| + \|\phi_n g - \int f g\| \leq \|F\| \|f - \phi_n g\| + \|g\| \|f - \phi_n\| \to 0\), the last term by Hölder.

Claim 5: \(g\) is unique a.e. (even if \(\mu\) not finite).

Let \(g_1\) and \(g_2\) be two such \(g\)'s. Then for \(f \in L^p\), \(\int f (g_1 - g_2) = 0\). For any \(E\) with \(\mu(E) < \infty\), \(f = \text{sgn}(g_1 - g_2) \chi_E \in L^p\). Then \(\int_E |g_1 - g_2| = 0\), so \(g_1 = g_2\) a.e. on \(E\). But \(\{x: g_1(x) \neq 0\text{ or } g_2(x) \neq 0\}\) is \(\sigma\)-finite, so \(g_1 = g_2\) a.e.

Now assume \(\mu\) is \(\sigma\)-finite. Write \(X = \bigcup X_n\) with \(X_1 \subseteq X_2 \subseteq \ldots\) and \(\mu(X_n) < \infty\). By considering the finite measure \(\mu_n(A) = \mu(A \cap X_n)\), we can define \(g_n\) by \(F(f) = \int g_n f \, d\mu\) when \(\text{ supp } f \subseteq X_n\). W.l.o.g. \(g_n(x) = 0\) for \(x \notin X_n\), and \(g_n(x) = g_n(x)\) for all \(x \in X_n \cap X_m\) (by a.e. uniqueness of \(g_n\)). Note that \(\|g_n\|_q \leq \|F\| \|\chi_{X_n}\|_q \leq \|F\| \|\chi_{X_n}\|_q\), so if \(g(x) = \lim g_n(x)\) then \(\|g\|_q \leq \|F\|\) by MCT. If \(f \in L^p\) let \(f_n = f \chi_{X_n}\). Then \(|\int \phi_n f_n - \int f g| \leq \|F(f) - F(f_n)\| + \|F(f_n) - \phi_n f_n - \int f g\| \leq \|F\| \|f - f_n\|_p + \|g\| \|f - \phi_n f_n\| \to 0\). But \(f_n \to f\), so \(\|f_n - f\|_p \to 0\) by DCT (dominate by \(|f|\)), and \(\int \phi_n f_n \to \int f g\) by DCT (dominate by \(|g|\) and use Hölder). Thus \(F(f) = \int f g\) and \(\|g\|_q \leq \|F\|\).

Now assume \(\mu\) is arbitrary but \(p > 1\), so \(q < \infty\). For all \(\sigma\)-finite \(E\), define \(g_E\) so that \(F(f) = \int f g_E\) when \(\text{ supp } f \subseteq E\). W.l.o.g. \(g_E(x) = 0\) when \(x \notin E\). Now \(\|g_E\|_q \leq \|F\| \|\chi_{L^p(E)}\|_q \leq \|F\| \|\chi_{L^p(X)}\|_q\), and if \(E \subseteq E'\) then \(\|g_{E'}\|_q \leq \|g_E\|_q\) since \(g_E = g_{E'}\) a.e. on \(E\). Thus we can choose an increasing sequence \(E_1 \subseteq E_2 \subseteq \ldots\) with \(\|g_{E_n}\|_q \to \lim_{E_n} \|g_{E_n}\|_q\). Let \(E = \bigcup E_i\) and suppose \(f \in L^p\). If \(F(f \chi_{E'}) \neq 0\) then since \(\text{ supp } f \) is \(\sigma\)-finite, there exists an \(E \subseteq X \setminus E\) with \(\mu(F) < \infty\) and \(F(f \chi_{E'}) \neq 0\). Thus \(\|g_{E_n}\|_q > 0\). But \(\|g_{E_n \cup F}\|_q = \|g_E\|_q + \|g_{E'}\|_q > \|g_{E_n}\|_q\), a contradiction. Hence \(F(f) = F(f \chi_{E_n}) = \int f g(E) d\mu\) for all \(f \in L^p\).

Finally, \(|F(f)| = \int f g| \leq \|f\|_p \|g\|_q\), so \(\|F\| \leq \|g\|_q\), and thus \(\|F\| = \|g\|_q\).
Lemma 1. Let \((X,d)\) be a metric space and \(\mu^*\) an outer measure such that \(\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)\) when \(d(A,B) > 0\). Then all Borel sets in \(X\) are \(\mu^*\)-measurable.

**Proof.** We show closed sets are measurable. We need \(\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \setminus A)\) for any \(E\) and any closed \(A\). W.l.o.g. \(\mu^*(E) < \infty\). Since \(A\) is closed, \(A = \{x : d(x,A) = 0\}\). Let \(A_n = \{x : d(x,A) < \varepsilon\}\). Let \(R_n = \{x \in E : \frac{1}{n+1} < d(x,A) \leq \frac{1}{n}\}\). Then \(d(R_n, R_m) > 0\) when \(|n - m| \geq 2\). Hence \(\sum_{n=1}^{N} \mu^*(R_{2n}) = \mu^*(\bigcup_{n=1}^{N} R_{2n}) \leq \mu^*(E) < \infty\), so \(\sum_{n=1}^{\infty} \mu^*(R_{2n})\) converges. Similarly \(\sum_{i=1}^{\infty} \mu^*(R_{2n+1})\) converges. Fix \(\varepsilon > 0\). Then for some \(N\), \(\sum_{n=N}^{\infty} \mu^*(R_n) < \varepsilon\). But \(E \setminus A = (E \setminus A_{1/N}) \cup \bigcup_{n=N}^{\infty} R_n\), so \(\mu^*(E \setminus A) \leq \mu^*(E \setminus A_N) + \varepsilon\) by countable subadditivity. Hence \(\mu^*(E \cap A) + \mu^*(E \setminus A) \leq \mu^*(E \cap A) + \mu^*(E \setminus A_{1/N}) + \varepsilon \leq \mu^*((E \cap A) \cup (E \setminus A_{1/N})) + \varepsilon \leq \mu^*(E) + \varepsilon\) since \(d(E \cap A, E \setminus A_{1/N}) \geq \frac{1}{N}\). Now let \(\varepsilon \to 0\). \qed

For \(\alpha > 0\) define \(m_{\alpha}^{(\varepsilon)}(A) = \inf \sum_{i=1}^{\infty} r_i^{\alpha}\) where the infimum is over all collections of balls \(B_{r_i}(x_i)\) with \(A \subseteq \bigcup_{i=1}^{\infty} B_{r_i}(x_i)\) and \(r_i \leq \varepsilon\). Define \(\mu^*(A) = \lim_{\varepsilon \to 0} m_{\alpha}^{(\varepsilon)}(A)\).

Lemma 2. For any metric space \((X,d)\), \(\mu^*\) exists, is an outer measure, and \(\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)\) when \(d(A,B) > 0\).

**Proof.** First note that \(m_{\alpha}^{(\varepsilon)}\) increases as \(\varepsilon\) decreases, so \(\mu^* = \lim_{\varepsilon} m_{\alpha}^{(\varepsilon)} = \lim_{\varepsilon} m_{\alpha}^{(1/n)}\) exists. The functions \(m_{\alpha}^{(\varepsilon)}\) are monotonic and countably subadditive: if \(A = \bigcup A_i\), choose \(B_{r_{ij}}(x_{ij})\) so that \(\sum_j r_{ij}^{\alpha} < m_{\alpha}^{(\varepsilon)} + \delta/2^i\). Then \(m_{\alpha}^{(\varepsilon)}(A) \leq \sum_{i,j} r_{ij}^{\alpha} = \sum m_{\alpha}^{(\varepsilon)}(A_i) + \delta\). Hence \(\mu^*\) is monotonic and countably subadditive: \(\mu^*(A) = \lim_{\varepsilon} \mu_{\alpha}^{(1/n)}(A) \leq \lim_{\varepsilon} \sum \mu_{\alpha}^{(1/n)}(A_i) = \sum \lim_{\varepsilon} \mu_{\alpha}^{(1/n)}(A_i) = \sum \mu^*(A_i)\) by discrete MCT. Finally, if \(\varepsilon < d(A,B)/2\) then \(m_{\alpha}^{(\varepsilon)}(A \cup B) = m_{\alpha}^{(\varepsilon)}(A) + m_{\alpha}^{(\varepsilon)}(B)\), so if \(d(A,B) > 0\) then \(\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)\). \qed

**Definition** The Borel measure \(\mu_\alpha\) that arises from \(\mu^*_\alpha\) is called the Hausdorff measure of dimension \(\alpha\). The Hausdorff dimension of a set \(A\) is \(\text{dim} A = \sup\{\alpha : \mu_\alpha(A) > 0\}\).

Lemma 3. If \(\alpha < \text{dim} A\) then \(\mu_\alpha(A) = \infty\).

**Proof.** If \(\alpha < \beta\) then \(m_{\beta}^{(\varepsilon)}(A) \leq \varepsilon^{\beta-\alpha} m_{\alpha}^{(\varepsilon)}(A)\). Thus if \(\mu_\beta(A) > 0\) then \(m_{\alpha}^{(\varepsilon)}(A) \geq \varepsilon^{\alpha-\beta} \mu_\beta(A) \to \infty\) as \(\varepsilon \to 0\). \qed

**Exercises**

1. Show that \(\mu_n\) is (up to a constant factor) the Lebesgue measure on \(\mathbb{R}^n\).

2. Show that the Cantor set has Hausdorff dimension \(\frac{\log 2}{\log 3}\).
Definition Let $\mathcal{K}$ be the set of compact subsets of a Hausdorff topological space $X$. A content on $X$ is a function $\lambda: \mathcal{K} \to [0, \infty]$ which is

1. monotone: $K_1 \subseteq K_2 \Rightarrow \lambda(K_1) \leq \lambda(K_2)$;
2. finitely additive: $K_1 \cap K_2 = \emptyset \Rightarrow \lambda(K_1 \cup K_2) = \lambda(K_1) + \lambda(K_2)$; and
3. finitely subadditive: $\lambda(K_1 \cup K_2) = \lambda(K_1) + \lambda(K_2)$ for any $K_1, K_2 \in \mathcal{K}$.

Lemma If $X$ is Hausdorff and $A$ and $B$ are disjoint compact sets then there exist disjoint open sets $U \supseteq A$ and $V \supseteq B$.

Proof. Fix $x \in A$. Then for all $y \in B$, there exists disjoint open $U_y$, $V_y$ with $x \in U_y$, $y \in V_y$. The $V_y$ cover $B$, so a finite collection $V_y$, do. Then $U = \bigcap U_y$, and $V = \bigcup V_y$, are disjoint open sets with $x \in U$ and $B \subseteq V$. Now repeat this process with each $x$ to get such sets $U(x)$ and $V(x)$. Since the $U(x)$ cover $A$, a finite subcollection do. Then $U = \bigcup U(x)$ and $V = \bigcap V(x)$ are as required. \qed

Lemma A content $\lambda$ gives rise to a measure $\mu$ on $(X, \sigma(\mathcal{K}))$ with $\mu(\mathcal{K}) \leq \lambda(K) \leq \mu(K)$.

Proof. Define the inner content of an open set $U$ by $\lambda_*(U) = \sup_{K \subseteq U} \lambda(K)$.
Define for any set $A$, $\mu^*(A) = \inf_{U \supseteq A} \lambda_*(U)$.
Use $K$, $K_i$ etc., to denote compact sets and $U$, $U_i$, etc., to denote open sets.

Both $\lambda_*$ and $\mu^*$ are clearly monotone. Suppose $K \subseteq U_1 \cup U_2$. Then $K \setminus U_1$ and $K \setminus U_2$ are disjoint compact sets, and so there are disjoint open $V_i \supseteq K \setminus U_i$. Then $K_i = K \setminus V_i$ are compact, $K_i \subseteq U_i$ and $K_1 \cup K_2 = K$. By induction, if $K \subseteq \bigcup_{i=1}^N U_i$, then there exists compact $K_i \subseteq U_i$ with $\bigcup_{i=1}^N K_i = K$. Now suppose $K \subseteq \bigcup_{i=1}^\infty U_i$. By compactness, $K \subseteq \bigcup_{i=1}^N U_i$ for some $N$, so we have $K_i \subseteq U_i$, $K = \bigcup K_i$, $i \leq N$, and $\lambda(K) \leq \sum_{i=1}^N \lambda(K_i) \leq \sum_{i=1}^\infty \lambda_*(U_i)$. Taking suprema over $K \subseteq U = \bigcup U_i$, $\lambda_*(U) \leq \sum \lambda_*(U_i)$ and so $\lambda_*$ is countably subadditive. Countable subadditivity of $\mu^*$ follows. Hence $\mu^*$ is an outer measure.

Fix any $E$ and pick $U \supseteq E$. Then

$$\lambda_*(U) \geq \sup_{K' \subseteq U \setminus K, K'' \subseteq U \setminus K'} \lambda(K' \cup K'') \quad K' \cup K'' \subseteq U$$

$$\geq \sup_{K' \subseteq U \setminus K, K'' \subseteq U \setminus K'} (\lambda(K') + \lambda(K'')) \quad K' \cap K'' = \emptyset$$

$$\geq \lambda_*(U \setminus K') + \lambda_*(U \setminus K') \quad U \setminus K' \text{ is open, definition of } \lambda_*$$

$$\geq \mu^*(E \setminus K') + \mu^*(E \setminus K) \quad E \cap K \subseteq U \setminus K', \text{ definition of } \mu^*$$

Taking infimums over $U$ we get $\mu^*(E) \geq \mu^*(E \setminus K) + \mu^*(E \cap K)$, so $K$ is measurable.

Finally, $\mu^*(U) = \inf_{U \supseteq U} \lambda_*(U') = \lambda_*(U)$, so $\mu^*(\mathcal{K}) = \lambda_*(\mathcal{K}) = \sup_{K' \subseteq \mathcal{K}} \lambda(K') \leq \lambda(K)$. $\mathcal{K}$ is measurable since it is the difference of two compact sets $K$ and $K \setminus \mathcal{K}$. Also, $\mu^*(K) = \inf_{U \supseteq K} \sup_{K' \subseteq U} \lambda(K') \geq \inf_U \lambda(K) = \lambda(K)$. \qed
Math 7351 15. Haar Measure Spring 2005

Definition  A topological group is a topological space $G$ which is also a group. Moreover, both the multiplication $\times : G \times G \to G$ and the inverse $(\cdot)^{-1} : G \to G$ are continuous (in the case of $\times$, continuity is with respect to the product topology on $G \times G$).

Examples
1. $\mathbb{R}$ under $\cdot$. More generally $\mathbb{R}^n$ with vector addition.
2. $\mathbb{R} \setminus \{0\}$ under $\cdot$. More generally the general linear group $GL_n(\mathbb{R})$ of all invertible $n \times n$ matrices with entries in $\mathbb{R}$. Multiplication is matrix multiplication. Topology can be given by considering $GL_n(\mathbb{R})$ as a subset of $\mathbb{R}^{n^2}$.
3. The special linear group $SL_n(\mathbb{R})$ (matrices of determinant 1) and the orthogonal group $O_n(\mathbb{R})$ (matrices $A$ with $AA^T = I$) are topological subgroups of $GL_n(\mathbb{R})$.

Lemma If $U$ is an open neighborhood of 1 in a topological group $G$ then there exists an open neighborhood $V$ of 1 such that $V^{-1}V = \{x^{-1}y : x, y \in V\} \subseteq U$.

Proof. By continuity of the map $(x, y) \mapsto x^{-1}y$, there exists a $V_1 \times V_2$ containing $(1, 1)$ with $V_1^{-1}V_2 \subseteq U$. Take $V = V_1 \cap V_2$. $\square$

Definition  A left Haar measure is a measure $\mu$ on the Borel sets of $G$ such that
1. $\mu$ is left invariant: if $g \in G$ then $\mu(gA) = \mu(A)$.
2. $\mu$ is outer regular: $\mu(A) = \inf \{\mu(U) : \text{open } U \supseteq A\}$.
3. If $K$ is compact then $\mu(K) < \infty$.
4. If $U$ is open then $\mu(U) > 0$.

Note $\int \chi_A(gx) \, d\mu(x) = \int \chi_{g^{-1}A}(x) \, d\mu(x) = \mu(g^{-1}A) = \mu(A) = \int \chi_A(x) \, d\mu$, so by standard arguments $\int f(gx) \, d\mu(x) = \int f(x) \, d\mu(x)$ for any integrable $f$.

Examples
1. Lebesgue measure on $(\mathbb{R}^n, +)$ is both a left and a right Haar measure.
2. The measure $\mu(E) = \int_E \frac{dx}{|x|}$ on $(\mathbb{R} \setminus \{0\}, \cdot)$ is a Haar measure.

Definition  A subset of $G$ is $\sigma$-bounded if it can be covered by a countable union of compact sets.

Exercises
1. If $K$ is compact then there exists a compact $G_\delta$-set $K'$ with $K \subseteq K'$, $\mu(K) = \mu(K')$. $[\mu(K) = \inf \{\mu(U) : U \supseteq K\}$, choose $U_i \supseteq K$ with $\mu(U_i) \to \mu(K)$ and inductively define $K_i+1 \subseteq K_i \cap U_i$ with $K \subseteq K_i+1$. Then $K' = \bigcap K_i \supseteq \bigcap K_i$.
2. If $K$ is a compact $G_\delta$-set and $G$ is $\sigma$-bounded then $\{(x, y) : xy \in K\}$ is measurable in $G \times G$. [Show the compact set $\{(x, y) : xy \in K\} \cap (K' \times K')$ is measurable first.]
**Definition** For any two sets $A$ and $B$, define $[A:B]$ to be the minimum $n$ such that $A$ can be covered with $n$ left translates of $B$, $A = \bigcup_{i=1}^{n} g_i B$. Note that if $K$ is compact and $U$ has non-empty interior then $[K:U] < \infty$.

**Theorem** If $G$ is a $\sigma$-bounded locally compact Hausdorff topological group, then there exists a left Haar measure on $G$. Moreover, it is unique up to multiplication by a positive constant.

**Proof.** Let $\mathcal{K}$ be the set of compact subsets of $G$. If $U$ is any open set then $K \cap U = K \setminus (K \setminus U)$ is a difference of compact sets, so lies in $\sigma(\mathcal{K})$. Since $G$ is $\sigma$-bounded, $G = \bigcup_{i=1}^{\infty} K_i$, so $U = \bigcup_{i=1}^{\infty} (K_i \cap U) \in \sigma(\mathcal{K})$. Thus $\sigma(\mathcal{K})$ contains all Borel sets.

By local compactness, there is a compact set $K_0$ with $1 \in K_0^0$. Define for all open $U \ni 1$ the function $\lambda_U(K) = [K:U]/[K_0:U]$. Note that $\lambda_U$ is subadditive and monotone, but not necessarily additive. Now $[K:U] \leq [K:K_0][K_0:U]$, so $0 \leq \lambda_U(K) \leq [K:K_0] < \infty$ for all $K \in \mathcal{K}$. Consider $\lambda_U$ as an element of the space $P = \prod_{K \in \mathcal{K}} [0,[K:K_0]]$ which is compact by Tychonoff. For each $U \ni 1$, let $S_U \subseteq P$ be the closure of $\{\lambda_U : 1 \in V \subseteq U\}$. Clearly each $S_U$ is closed and any finite intersection $\bigcap_{U \ni 1} S_U$ is non-empty since it contains $\lambda_{K_0}$. Hence $\bigcap_{U \ni 1} S_U \neq \emptyset$. Let $\lambda \in \bigcap_{U \ni 1} S_U$. Suppose $K$ and $K'$ are disjoint compact sets. Then $1 \notin K^{-1}K'$, and $K^{-1}K'$ is compact (continuous image of $K \times K'$), so closed. Thus there is an open $U \ni 1$ with $U^{-1}U \cap K^{-1}K' = \emptyset$, so $KU^{-1} \cap K'U^{-1} = \emptyset$. In this case $[K \cup K':V] = [K:V] + [K':V]$ for all $V \subseteq U$. Hence $\lambda(K \cup K') = \lambda(K) + \lambda(K')$ for all $\lambda \in S_U$ (the set of such $\lambda$ is closed and contains all $\lambda_V$, $V \subseteq U$). In particular $\lambda(K \cup K') = \lambda(K) + \lambda(K')$. Hence $\lambda$ is a content on $G$ and gives rise to a measure $\mu$ on $(G, \sigma(\mathcal{K}))$. Now $\lambda_U(gK) = \lambda_U(K)$ for all $U$, $K$, and $g$, so $\lambda(gK) = \lambda(K)$. Thus $\mu(gA) = \mu(A)$ for all $A \in \sigma(\mathcal{K})$. Finally, $\lambda \in P$, so $\lambda(K) \leq [K:K_0] < \infty$. Now $K' = KK_0$ is compact and $K \subseteq K'$, so $\mu(K) \leq \mu(K') \leq \lambda(K') < \infty$. Clearly $\lambda(K_0) = 1$, and if $U$ is non-empty and open, $K_0 \subseteq \bigcup_{i=1}^{\infty} g_i K_0$, so $1 \leq n \mu(U)$, and $\mu(U) > 0$.

**Uniqueness.** Fix $A \in \mathcal{K}$ with non-empty interior, so $0 < \mu(A), \nu(A) < \infty$. Define $c(g) = \mu(A)/\mu(Ag^{-1})$, so that $\mu(A) = \int \chi_A(xg)c(g) \, d\mu(x)$. Note $\mu(Ag^{-1}) = \int \chi_{gA}(x(g)) \, d\nu(x)$ so is measurable as a function of $g$ (Fubini). Then for any $B \in \mathcal{K}_0$

$$
\mu(A)\nu(B) = \int \int \chi_A(xy)c(y)\chi_B(y) \, d\mu(x) \, d\nu(y) \\
= \int \chi_A(y)\chi_B(x^{-1}y) \, d\nu(y) \, d\mu(x) \\
= \int \chi_A(y)c((y^{-1}x^{-1})\chi_B((y^{-1}x^{-1})) \, d\mu(x) \, d\nu(y) \\
= \nu(A) \int c(x^{-1})\chi_B(x^{-1}) \, d\mu(x).
$$

(All functions measurable in $G \times G$ when $A, B \in \mathcal{K}_0$). Applying the same argument with $\nu$ replaced by $\mu$ gives $\int c(x^{-1})\chi_B(x^{-1}) \, d\mu(x) = \mu(B)$, so $\mu(A)\nu(B) = \nu(A)\mu(B)$ and if $\alpha = \nu(A)/\mu(A)$ then $\mu(B) = \alpha \nu(B)$ for all $B \in \mathcal{K}_0$. Pick $K', K'' \in \mathcal{K}_0$ so that $K \subseteq K', K''$ and $\mu(K) = \mu(K')$, $\nu(K) = \nu(K'')$, then $\mu(K) = \mu(K' \cap K'') = \nu(K' \cap K'') = \nu(K)$, so $\mu = \alpha \nu$ on $K$ and hence on $\sigma(\mathcal{K})$. □