Extension of measures through integration

Bengt Ringnér∗†

November 7, 2008

1 Measure

Definition 1 A nonempty set, \(\mathcal{R}\), of subsets of a set, \(\mathcal{U}\), is called a ring if \(A, B \in \mathcal{R}\) implies \(A \cup B, A \cap B, A \cap B^c \in \mathcal{R}\).

It is called an algebra if, in addition, \(\mathcal{U} \in \mathcal{R}\).

An algebra, \(\mathcal{A}\), is called a \(\sigma\)-algebra if \(A_1, A_2, \ldots \in \mathcal{A}\) implies \(\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}\).

For instance, all bounded intervals of the real line together with all finite unions of them form a ring.

Definition 2 A (nonnegative) measure, \(\mu\), is a function, defined on a ring, attaining nonnegative real values or infinity, and such that,

1. if \(A\) and \(B\) are disjoint, then \(\mu(A \cup B) = \mu(A) + \mu(B)\), and
2. if \(A_n \not\to \emptyset \in \mathcal{R}\) as \(n \to \infty\), that is, \(A_1 \subset A_2 \subset \ldots\), and \(A = \bigcup_{n=1}^{\infty} A_n \in \mathcal{R}\), then \(\mu(A_n) \to \mu(A)\) as \(n \to \infty\).

If the ring is an algebra and \(\mu(\mathcal{U}) = 1\), then \(\mu\) is called a probability measure.

Remark Since \(A_n \not\to \emptyset \Rightarrow A_1 \cap A_n^c \not\to A_1\) and \(\mu(A_1 \cap A_n^c) + \mu(A_n) = \mu(A_1)\), it follows that:

- If \(\mu(A_1) < \infty\), and \(A_n \not\to \emptyset\) as \(n \to \infty\), then \(\mu(A_n) \to 0\).

In most applications \(\mu(A) < \infty\) for all \(A \in \mathcal{R}\). In this case 2. in the definition can be replaced by •. This follows from \(A_n \not\to A \Rightarrow A \cap A_n^c \not\to \emptyset\) and \(\mu(A \cap A_n^c) + \mu(A_n) = \mu(A)\).

∗Centre for Mathematical Sciences, Lund University, Lund, Sweden.
†Homepage: http://www.maths.lth.se/~bengtr
A theorem of central importance, both in integration theory and in probability says that the domain of definition of any measure can be extended to a $\sigma$-algebra. Traditionally, the theorem is proved by introducing an outer measure by

$$\mu^*(E) = \inf \sum_{k=1}^{\infty} \mu(A_k)$$

where the infimum extends over all possibilities such that $E \subset \bigcup_{n=1}^{\infty} A_k$ with the $A_k$ in $\mathcal{R}$. Then one shows that $\mu^*$ is a measure when it is restricted to the smallest $\sigma$-algebra containing $\mathcal{R}$. This is done by defining what is meant by a measurable set, showing that these form a $\sigma$-algebra containing $\mathcal{R}$, that $\mu^*$ restricted to the measurable sets is a a measure, and that $\mu^* = \mu$ on $\mathcal{A}$. The details, which become rather awkward, can be found in most textbooks.

Instead of working with sets, we shall use their indicator functions, defined by

$$1_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases},$$

and work with integrals and the relation $\mu(A) = I(1_A)$. In the next section the integral is extended with the aid of an upper integral similar to the outer measure.

### 1.1 Integration, first part

A general procedure to build integration theory is to start with a suitable class of “nice” functions, define an integral with respect to $\mu$, and extend it to a larger class of functions. The “nice” functions used here are step functions, namely functions that can be written as

$$\varphi(x) = \sum_{i=1}^{n} a_k 1_{A_k}$$

with $a_k$ real constants, $A_k$ in the ring $\mathcal{R}$, and $\mu(A_k) < \infty$ for all $k$. Define the integral of such a function by

$$I(\varphi) = \sum_y y\mu(\{x : \varphi(x) = y\}) = \sum_{i=1}^{n} a_k \mu(A_k).$$

As in the case of the Riemann integral, one can convince oneself that the two expressions are equal.
In Riemann theory, one estimates a general integrand, \( f \), from below and above with step functions. The method used here is to approximate \( f \) such that the integral of \( |f - \varphi| \) can be made arbitrarily small in a sense to be described in a moment. This method seems to be first published by Bourbaki [1], Chap. 4.3.4, Def. 2 and Prop. 7 pp. 129, 130, when \( \Omega \) is a locally compact topological space. According to Hörmander, at the end of [2], this restriction is unessential. To give the approximation a meaning, define the upper integral of a nonnegative function, \( f \), by

\[
I^*(f) = \inf \sum_{k=1}^{\infty} I(\psi_k)
\]

where the infimum extends over all possibilities where the \( \psi_k \) are nonnegative “nice” functions with \( f(x) \leq \sum_{k=1}^{\infty} \psi_k(x) \) for all \( x \). Since \( \sum \psi_n(x) \) can be infinite, \( f \) is allowed to attain infinity. The upper integral might nevertheless be finite.

**Lemma 1** The upper integral, \( I^* \), satisfies

1. \( I^*(0) = 0 \),
2. \( I^*(f) \leq I^*(g) \) if \( f \leq g \) and,
3. \( I^*(\sum_{n=1}^{\infty} f_n) \leq \sum_{n=1}^{\infty} I^*(f_n) \).

In 3. \( \sum_{n=1}^{\infty} f_n \) can denote any function given by

\[
(\sum_{n=1}^{\infty} f_n)(x) = \begin{cases} 
\sum_{n=1}^{\infty} f_n(x) & \text{if } \sum_{n=1}^{\infty} f_n(x) < \infty \\
\alpha & \text{otherwise},
\end{cases}
\]

where \( 0 \leq \alpha \leq \infty \) is arbitrary.

**Proof** The first two assertions are obvious. The third one follows by choosing \( \psi_{n,k} \) such that \( \sum_{k=1}^{\infty} I(\psi_{n,k}) < I^*(f_n) + \varepsilon/2^n, n = 1, 2, \ldots \). □

A function, \( f \), which may attain values in \([-\infty, \infty]\), is called integrable if there exist “nice” functions \( \varphi_1, \varphi_2, \ldots \) such that

\[
I^*(|f - \varphi_n|) \to 0
\]
as \( n \to \infty \). To be able to define

\[
I(f) = \lim_{n \to \infty} I(\varphi_n)
\]
we have to show that
• the limit exists, and
• it is independent of the choice of approximating sequence.

This is done with the Cauchy convergence criterion. First use

\[ |I(\varphi_m) - I(\varphi_n)| \leq I(|\varphi_m - \varphi_n|). \]

From Corollary 1 to the lemma below we get

\[ I(|\varphi_m - \varphi_n|) \leq I^*(|\varphi_m - \varphi_n|). \]

Finally

\[ I^*(|\varphi_m - \varphi_n|) = I^*(|\varphi_m - f - (\varphi_n - f)|) \]
\[ \leq I^*(|\varphi_m - f| + |\varphi_n - f|) \]
\[ \leq I^*(|\varphi_m - f|) + I^*(|\varphi_n - f|) \rightarrow 0 \]

as \( m, n \rightarrow \infty \). Therefore, the limit exists. Repeating the same argument with the difference, \( \varphi_n' - \varphi_n'' \), between two choices of \( \varphi \)-sequences, gives that the limit is independent of the choice of the \( \varphi \)-sequence.

The lemma we need corresponds to the condition for \( \mu \) in the remark after Definition 2.

**Lemma 2** If \( \varphi_1, \varphi_2, \ldots \) are step functions as above, and \( \varphi_n \searrow 0 \) pointwise as \( n \rightarrow \infty \), then

\[ I(\varphi_n) \rightarrow 0 \]

as \( n \rightarrow \infty \).

**Proof** For a given \( \varepsilon > 0 \) let

\[ A_n = \{ x : \varphi_n(x) \geq \varepsilon \} \in \mathcal{R}. \]

Then \( A_n \searrow \emptyset \) and \( \mu(A_1) < \infty \), which implies \( \mu(A_n) \rightarrow 0 \). Therefore, from

\[ I(\varphi_n) \leq \varepsilon \mu(\{ x : \varphi_1 > 0 \}) + \mu(A_n) \max_x \varphi_1(x), \]

it is clear that \( I(\varphi_n) \rightarrow 0 \). \( \square \)

As a corollary we get the counterpart of \( \mu^* = \mu \) on \( \mathcal{R} \).

**Corollary 1** If \( \varphi, \psi_1, \psi_2, \ldots \) are nonnegative step functions, and \( \varphi \leq \sum_{k=1}^{\infty} \psi_k \), then

\[ I(\varphi) \leq \sum_{k=1}^{\infty} I(\psi_k). \]

Consequently, \( I^*(\varphi) = I(\varphi) \).
Proof

\[ I(\varphi) - \sum_{k=1}^{\infty} I(\psi_k) \leq I\left(\varphi - \sum_{k=1}^{n} \psi_k\right) \leq I\left(\max(0, \varphi - \sum_{k=1}^{n} \psi_k)\right) \rightarrow 0 \]

as \( n \to \infty. \)

**Theorem 1** If \( f_1 \) and \( f_2 \) are integrable, then so are all linear combinations of them as well as \( \max(f_1, f_2) \) and \( \min(f_1, f_2) \). It holds

\[ I(a_1 f_1 + a_2 f_2) = a_1 I(f_1) + a_2 I(f_2), \]

and

\[ f_1 \leq f_2 \Rightarrow I(f_1) \leq I(f_2). \]

Consequently, \( f \) is integrable if and only if \( f^+ \), (\( = \max(0, f) \)), and \( f^- \), (\( = \max(0, -f) \)), are integrable. Further \( |f| = \max(f, -f) \) is integrable if \( f \) is so.

Proof Linearity follows from \( |a_1 f_1 + a_2 f_2 - (a_1 \varphi_1 + a_2 \varphi_2)| \leq |a_1||f_1 - \varphi_1| + |a_2||f_2 - \varphi_2| \), and the second formula from \( |f - \varphi| \leq |f - \max(0, \varphi)| \) when \( f \geq 0 \). The rest follows from \( ||f| - |\varphi|| \leq |f - \varphi|, \) \( \max(f_1, f_2) = (f_1 + f_2 + |f_1 - f_2|)/2 \), and a similar formula for \( \min \).

**Theorem 2** (Beppo Levi) If \( f_1, f_2, \ldots \) are nonnegative and integrable with \( \sum_{n=1}^{\infty} I(f_n) < \infty \), then \( \sum_{n=1}^{\infty} f_n \) is integrable, and

\[ I\left(\sum_{n=1}^{\infty} f_n\right) = \sum_{n=1}^{\infty} I(f_n). \]

Conversely, if \( \sum_{n=1}^{\infty} f_n \) is integrable, then \( \sum_{n=1}^{\infty} I(f_n) < \infty \).

Proof For a given \( \varepsilon > 0 \) choose “nice” functions \( \varphi_1, \ldots \) so that \( I^*(|f_n - \varphi_n|) < \varepsilon/2^n \), and \( |I(f_n) - I(\varphi_n)| < \varepsilon/2^n \) (in fact, this follows automatically). Now

\[
I^*(\left|\sum_{n=1}^{\infty} f_n - \sum_{n=1}^{N} \varphi_n\right|) \leq \sum_{n=1}^{N} I^*(|f_n - \varphi_n|) + \sum_{n=N+1}^{\infty} I^*(f_n) \\
\leq \varepsilon + \sum_{n=N+1}^{\infty} I^*(f_n) < 2\varepsilon
\]
if \( N \) is large enough, provided we can show that \( \sum_{n=1}^{\infty} I^*(f_n) < \infty \). As in the proof of the preceding theorem, we may assume \( \varphi_n \geq 0 \). Hence,

\[
I^*(f_n) \leq I^*(\varphi_n) + I^*(|f_n - \varphi_n|) \\
\leq I(\varphi_n) + I^*(|f_n - \varphi_n|) \\
\leq I(f_n) + \varepsilon/2^n + \varepsilon/2^n.
\]

This proves the first part. The converse follows from

\[
\sum_{n=1}^{N} I(f_n) = I\left( \sum_{n=1}^{N} f_n \right) \leq I\left( \sum_{n=1}^{\infty} f_n \right)
\]

for all \( N \).

**Remark** The integral does not depend how \( \sum f_n \) is defined when \( \sum f_n(x) = \infty \). This means that we can compute integrals like

\[
\int_{-1}^{1} \frac{1}{\sqrt{|x|}} \, dx = \lim_{n \to \infty} \int_{-1}^{1} \min(n, \frac{1}{\sqrt{|x|}}) \, dx = 4.
\]

**Corollary 2** (Monotone Convergence) If \( f_1 \leq f_2 \leq \ldots \) or \( f_1 \geq f_2 \geq \ldots \) are integrable, then

\[
I(\lim_{n \to \infty} f_n) = \lim_{n \to \infty} I(f_n)
\]

in the sense that the limit is integrable if and only if the limit on the right is finite.

**Proof** Follows from \( f_n = f_1 + \sum_{k=1}^{n-1} (f_{k+1} - f_k) \).

**Corollary 3** (Dominated Convergence, Lebesgue) If \( f_1, \ldots, g \) are integrable, \( |f_n| \leq g \), and \( f_n \to f \) as \( n \to \infty \), then \( f \) is integrable and

\[
I(f_n) \to I(f).
\]

**Proof** Let \( g_n = \sup(f_n, f_{n+1}, \ldots) \) and \( h_n = \max(f_n, f_{n+1}, \ldots, f_{n+k}) \). Then the \( h_n \) are integrable, \( h_n \not\uparrow g_n \) as \( k \to \infty \) and \( I(h_n) \leq I(g) \), so the \( g_n \) are integrable. Next \( g_n \not\downarrow f \) with \( I(g_n) \geq -I(g) \), so \( f \) is integrable, and

\[
\limsup_{n \to \infty} I(f_n) \leq \lim_{n \to \infty} I(g_n) = I(f).
\]

Similar calculations for \( \liminf \) give the desired result.
1.2 Measurable functions

This section contains a digression and can be omitted.

**Definition 3** A function, $f$, is called measurable if $\min(g, \max(-g, f))$ is integrable for all integrable nonnegative $g$.

It follows directly from the definition that if $f$ is measurable and $|f| \leq g$ with $g$ integrable, then $f$ is integrable. Further, Dominated Convergence gives that all limits of measurable functions are again measurable.

**Theorem 3** If $f_1$ and $f_2$ are measurable, then so is $f_1 + f_2$.

**Proof** For a given integrable $g \geq 0$, let $f_{n}(x) = \min(n g, \max(-n g, f_{n}))$. Then $f_{n}(x) \to f(x)$ for all $x$ with $g(x) > 0$, so

$$\min(g, \max(-g, f_{1} + f_{2})) \to \min(g, \max(-g, f_{1} + f_{2}))$$

as $n \to \infty$. Since all $f_{n}$ are integrable, the left hand side is so, and Dominated Convergence completes the proof. \qed

1.3 Measure theory

One defines the integral over a set by

$$I_{A}(f) = I(1_{A}f) = I(1_{A}f^{+}) - I(1_{A}f^{-}).$$

If $A$ is such that $1_{A}$ is integrable, this is possible, since $\min(n 1_{A}, f) \to 1_{A}f$ as $n \to \infty$, and . But it is possible also for other sets, for instance $\mathcal{U}$. In the next theorem it is shown that those sets for which $I_{A}(f)$ is defined for all integrable $f$ constitute a $\sigma$-algebra.

**Theorem 4** Let $\mu$ be a measure defined on a ring $\mathcal{R}$ of subsets of a set $\mathcal{U}$. Then $\mu$ can be extended to a $\sigma$-algebra containing $\mathcal{R}$ without changing its values on $\mathcal{R}$.

**Proof** We shall call a set, $A$, measurable if $1_{A}f$ is integrable for all integrable $f \geq 0$, and define the extended measure for those sets by

$$\tilde{\mu}(A) = \begin{cases} I(1_{A}) & \text{if } 1_{A} \text{ is integrable} \\ \infty & \text{otherwise.} \end{cases}$$

As noted above, all sets with $1_{A}$ integrable are measurable. Then:
1. \( \mathcal{U} \) is measurable, since \( 1_{\mathcal{U}} f = f \).

2. If \( A \) is measurable, then so is \( A^c \), since \( 1_{A^c} f = f - 1_A f \).

3. (a) If \( A \) and \( B \) are measurable, then so are \( A \cup B \) and \( A \cap B \), since 
   \[
   1_{A \cup B} f = \max(1_A f, 1_B f), \quad \text{and} \quad 1_{A \cap B} f = \min(1_A f, 1_B f).
   \]
   (b) Further, if \( A \) and \( B \) are disjoint, then
   \[
   \tilde{\mu}(A \cup B) = \tilde{\mu}(A) + \tilde{\mu}(B).
   \]
   When both \( A \) and \( B \) have finite measures, this is clear. If \( A \cup B \)
   has finite measure, the choice of \( f = 1_{A \cup B} \) shows that all its
   measurable subsets have finite measures. Therefore, if the right
   hand side is infinite, then so is the left one.

4. (a) If \( A_1, A_2, \ldots \) are measurable, then so is \( \bigcup_{n=1}^{\infty} A_n \). This follows
   from \( 1_{\bigcup_{n=1}^{N} A_n} f / 1_{\bigcup_{n=1}^{\infty} A_n} f \) as \( N \to \infty \) and Monotone Conver-
   gence.
   (b) Further, if \( A_1 \subset A_2 \subset \ldots \), then
   \[
   \tilde{\mu}(A_n) / \tilde{\mu}(\bigcup_{n=1}^{\infty} A_n)
   \]
   as \( n \to \infty \). This follows from Monotone Convergence if the se-
   quence is bounded, while, otherwise, the previous reasoning shows
   that the right hand side is infinite.

This shows that \( \tilde{\mu} \) is a measure on the \( \sigma \)-algebra of measurable sets. It
remains to show that \( \tilde{\mu} \) is defined on \( \mathcal{R} \) and coincides with \( \mu \) there.

1. Every \( A \in \mathcal{R} \) is measurable since \( |1_A f - 1_A \varphi| \leq |f - \varphi| \).

2. If \( A \in \mathcal{R} \), then \( \tilde{\mu}(A) = \mu(A) \).
   (a) If \( \mu(A) < \infty \) this is clear.
   (b) If \( \mu \) is \( \sigma \)-finite, that is \( \mathcal{U} = \bigcup_{n=1}^{\infty} R_n \) for some \( R_n \in \mathcal{R} \) with \( \mu(R_n) < \infty \),
   then Definition 2 gives \( \mu(A) = \lim_{N \to \infty} \mu(\bigcup_{n=1}^{N} A \cap R_n) \), while
   4b above gives the same for \( \tilde{\mu} \).
   (c) This is more complicated, but true.

\( \square \)

8
2 The Lebesgue integral

Here we are faced with a measure $\mu$ defined on a $\sigma$-algebra of subsets of a set $U$ and want to integrate real valued functions defined on $U$.

**Definition 4** A function $f$ is called measurable if the sets 
\[
\{x : f(x) > y\}
\]
belong to the $\sigma$-algebra for all $y$.

**Definition 5** The Lebesgue integral of a nonnegative measurable $f$ is given by
\[
\int f \, d\mu = \lim_{j \to \infty} S_{2j},
\]
where
\[
S_n = \sum_{k=1}^{n^2-1} \frac{k}{n} \mu\left(\{x : \frac{k}{n} < f(x) \leq \frac{k+1}{n}\}\right) + n\mu\left(\{x : f(x) > n\}\right),
\]
provided the limit exists. The integral of a general measurable function is
\[
\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu,
\]
provided both integrals on the right hand side exist. The integral over a set in the $\sigma$-algebra is
\[
\int_A f \, d\mu = \int 1_A f \, d\mu.
\]

Since any $\sigma$-algebra is a ring, we can still use the integral introduced earlier and its properties.

**Theorem 5** For a measurable $f$,
\[
\int f \, d\mu = I(f)
\]
in the sense that if one side exists, then so does the other one.

**Proof**
\[
S_n = \sum_{k=1}^{n^2-1} \frac{k}{n} I\left(\{x : \frac{k}{n} < f(x) \leq \frac{k+1}{n}\}\right) + nI\left(\{x : f(x) > n\}\right)
\]
\[
= I\left(\sum_{k=1}^{n^2-1} \frac{1}{n} \{x : \frac{k}{n} < f(x) \leq \frac{k+1}{n}\}\right) + n1\{x : f(x) > n\}.
\]
But $f_{2j} \nrightarrow f$ as $j \to \infty$, so Monotone Convergence gives the result. □

3 Stone-Daniell representation

Definition 6 A vector space, $L$, of real-valued functions defined on a set $\mathcal{U}$ is called a lattice if

- $\varphi \in L$ implies $|\varphi| \in L$.

It is called a Stone lattice if, in addition

- $\varphi \in L$ implies $\min(1, \varphi) \in L$.

Definition 7 An integral is a nonnegative linear functional, $I$, defined on a lattice and such that

- $\varphi_n \downarrow 0$ pointwise implies $I(\varphi_n) \to 0$ as $n \to \infty$.

Theorem 6 (Stone 1948) If $I$ is an integral defined on a Stone lattice, then there exists a measure $\mu$ such that

$$I(\varphi) = \int \varphi \, d\mu.$$ 

Proof In the proofs above, just replace the “nice” step functions with $L$-functions to get the measure. It remains to prove that all $L$-functions are measurable, so that $\int \varphi \, d\mu$ is defined. This is where the Stone property comes in. Consider

$$n\left( \min(y + 1/n, \varphi) - \min(y, \varphi) \right) \to 1_{\{x: \varphi(x) > y\}}$$

as $n \to \infty$. For $y > 0$ the functions to the left are integrable. Since the left hand side is less than $\varphi/y$ for all $n$ and $I(\varphi/y) < \infty$, the right hand side is integrable, so the sets $\{x: \varphi(x) > y\}$ are measurable according to the definition in the theorem. For $y \leq 0$, use the fact that $\{x: \varphi(x) > -y\}$ and complements and limits of measurable sets are measurable. □

References
