2. Measurable functions and random variables

2.1. Measurable functions. Let \((E, \mathcal{E})\) and \((G, \mathcal{G})\) be measurable spaces. A function \(f : E \to G\) is measurable if \(f^{-1}(A) \in \mathcal{E}\) whenever \(A \in \mathcal{G}\). Here \(f^{-1}(A)\) denotes the inverse image of \(A\) by \(f\):

\[
f^{-1}(A) = \{x \in E : f(x) \in A\}.
\]

Usually \(G = \mathbb{R}\) or \(G = [-\infty, \infty]\), in which case \(\mathcal{G}\) is always taken to be the Borel \(\sigma\)-algebra. If \(E\) is a topological space and \(\mathcal{E} = \mathcal{B}(E)\), then a measurable function on \(E\) is called a Borel function. For any function \(f : E \to G\), the inverse image preserves set operations

\[
f^{-1}\left(\bigcup_i A_i\right) = \bigcup_i f^{-1}(A_i), \quad f^{-1}(G \setminus A) = E \setminus f^{-1}(A).
\]

Therefore, the set \(\{f^{-1}(A) : A \in \mathcal{G}\}\) is a \(\sigma\)-algebra on \(E\) and \(\{A \subseteq G : f^{-1}(A) \in \mathcal{E}\}\) is a \(\sigma\)-algebra on \(G\). In particular, if \(\mathcal{G} = \sigma(\mathcal{A})\) and \(f^{-1}(A) \in \mathcal{E}\) whenever \(A \in \mathcal{A}\), then \(\{A : f^{-1}(A) \in \mathcal{E}\}\) is a \(\sigma\)-algebra containing \(\mathcal{A}\) and hence \(\mathcal{G}\), so \(f\) is measurable. In the case \(G = \mathbb{R}\), the Borel \(\sigma\)-algebra is generated by intervals of the form \((-\infty, y], y \in \mathbb{R}\), so, to show that \(f : E \to \mathbb{R}\) is Borel measurable, it suffices to show that \(\{x \in E : f(x) \leq y\} \in \mathcal{E}\) for all \(y\).

If \(E\) is any topological space and \(f : E \to \mathbb{R}\) is continuous, then \(f^{-1}(U)\) is open in \(E\) and hence measurable, whenever \(U\) is open in \(\mathbb{R}\); the open sets \(U\) generate \(\mathcal{B}\), so any continuous function is measurable.

For \(A \subseteq E\), the indicator function \(1_A\) of \(A\) is the function \(1_A : E \to \{0, 1\}\) which takes the value 1 on \(A\) and 0 otherwise. Note that the indicator function of any measurable set is a measurable function. Also, the composition of measurable functions is measurable.

Given any family of functions \(f_i : E \to G, i \in I\), we can make them all measurable by taking

\[
\mathcal{E} = \sigma(f_i^{-1}(A) : A \in \mathcal{G}, i \in I).
\]

Then \(\mathcal{E}\) is the \(\sigma\)-algebra generated by \((f_i : i \in I)\).

**Proposition 2.1.1.** Let \(f_n : E \to \mathbb{R}, n \in \mathbb{N}\), be measurable functions. Then so are \(f_1 + f_2, f_1 f_2\) and each of the following:

\[
\inf_n f_n, \quad \sup_n f_n, \quad \liminf_n f_n, \quad \limsup_n f_n.
\]

**Theorem 2.1.2** (Monotone class theorem). Let \((E, \mathcal{E})\) be a measurable space and let \(\mathcal{A}\) be a \(\pi\)-system generating \(\mathcal{E}\). Let \(\mathcal{V}\) be a vector space of bounded functions \(f : E \to \mathbb{R}\) such that:

(i) \(1 \in \mathcal{V}\) and \(1_A \in \mathcal{V}\) for all \(A \in \mathcal{A}\);

(ii) if \(f_n \in \mathcal{V}\) for all \(n\) and \(f\) is bounded with \(0 \leq f_n \uparrow f\), then \(f \in \mathcal{V}\).

Then \(\mathcal{V}\) contains every bounded measurable function.
Proof. Consider \( \mathcal{D} = \{ A \in \mathcal{E} : 1_A \in \mathcal{V} \} \). Then \( \mathcal{D} \) is a \( \sigma \)-system containing \( \mathcal{A} \), so \( \mathcal{D} = \mathcal{E} \). Since \( \mathcal{V} \) is a vector space, it thus contains all finite linear combinations of indicator functions of measurable sets. If \( f \) is a bounded and non-negative measurable function, then the functions \( f_n = 2^{-n} \lfloor 2^n f \rfloor \), \( n \in \mathbb{N} \), belong to \( \mathcal{V} \) and \( 0 \leq f_n \uparrow f \), so \( f \in \mathcal{V} \). Finally, any bounded measurable function is the difference of two non-negative such functions, hence in \( \mathcal{V} \).

2.2. Image measures. Let \((E, \mathcal{E})\) and \((G, \mathcal{G})\) be measurable spaces and let \( \mu \) be a measure on \( \mathcal{E} \). Then any measurable function \( f : E \to G \) induces an image measure \( \nu = \mu \circ f^{-1} \) on \( \mathcal{G} \), given by

\[
\nu(A) = \mu(f^{-1}(A)).
\]

We shall construct some new measures from Lebesgue measure in this way.

Lemma 2.2.1. Let \( g : \mathbb{R} \to \mathbb{R} \) be non-constant, right-continuous and non-decreasing. Let \( I = (g(-\infty), g(\infty)) \) and define \( f : I \to \mathbb{R} \) by \( f(x) = \inf \{ y \in \mathbb{R} : x \leq g(y) \} \). Then \( f \) is left-continuous and non-decreasing. Moreover, for \( x \in I \) and \( y \in \mathbb{R} \),

\[
f(x) \leq y \quad \text{if and only if} \quad x \leq g(y).
\]

Proof. Fix \( x \in I \) and consider the set \( J_x = \{ y \in \mathbb{R} : x \leq g(y) \} \). Note that \( J_x \) is non-empty and is not the whole of \( \mathbb{R} \). Since \( g \) is non-decreasing, if \( y \in J_x \) and \( y' \geq y \), then \( y' \in J_x \). Since \( g \) is right-continuous, if \( y_n \in J_x \) and \( y_n \downarrow y \), then \( y \in J_x \). Hence \( J_x = [f(x), \infty) \) and \( x \leq g(y) \) if and only if \( f(x) \leq y \). For \( x \leq x' \), we have \( J_x \supseteq J_{x'} \) and so \( f(x) \leq f(x') \). For \( x_n \uparrow x \), we have \( J_x = \cap_n J_{x_n} \), so \( f(x_n) \to f(x) \). So \( f \) is left-continuous and non-decreasing, as claimed.

Theorem 2.2.2. Let \( g : \mathbb{R} \to \mathbb{R} \) be non-constant, right-continuous and non-decreasing. Then there exists a unique Radon measure \( dg \) on \( \mathbb{R} \) such that, for all \( a, b \in \mathbb{R} \) with \( a < b \),

\[
dg((a, b]) = g(b) - g(a).
\]

Moreover, we obtain in this way all non-zero Radon measures on \( \mathbb{R} \).

The measure \( dg \) is called the Lebesgue-Stieltjes measure associated with \( g \).

Proof. Define \( I \) and \( f \) as in the lemma and let \( \mu \) denote Lebesgue measure on \( I \). Then \( f \) is Borel measurable and the induced measure \( dg = \mu \circ f^{-1} \) on \( \mathbb{R} \) satisfies

\[
dg((a, b]) = \mu(\{ x : f(x) > a \text{ and } f(x) \leq b \}) = \mu((g(a), g(b)) = g(b) - g(a).
\]

The argument used for uniqueness of Lebesgue measure shows that there is at most one Borel measure with this property. Finally, if \( \nu \) is any Radon measure on \( \mathbb{R} \), we can define \( g : \mathbb{R} \to \mathbb{R} \), right-continuous and non-decreasing, by

\[
g(y) = \begin{cases} 
\nu((0, y]), & \text{if } y \geq 0, \\
-\nu((y, 0]], & \text{if } y < 0.
\end{cases}
\]

Then \( \nu((a, b]) = g(b) - g(a) \) whenever \( a < b \), so \( \nu = dg \) by uniqueness.

\[\square\]