Chapter 2

Measurable Functions

2.1 Prelude - Liminf and Limsup

This section has nothing to do with measure theory as such. It introduces some important tools from analysis which there wasn't time to cover in MAS207. Let (a_n) be a sequence of real numbers. It may or may not converge. For example the sequence whose nth term is $(-1)^n$ fails to converge but it does have two convergent subsequences corresponding to $a_{2n-1} = -1$ and $a_{2n} = 1$. This is a very special case of a general phenomenon that we'll now describe.

Assume that the sequence (a_n) is *bounded*, i.e. there exists $K > 0$ so that $|a_n| \leq K$ for all $n \in \mathbb{N}$. Define a new sequence (b_n) by $b_n = \inf_{k \geq n} a_k$. Then you can check that (b_n) is monotonic increasing and bounded above (by K). Hence it converges to a limit. We define $\liminf_{n\to\infty} a_n = \lim_{n\to\infty} b_n$ and call it the *limit inferior* or *liminf* (for short) of the sequence (a_n). So we have

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} \inf_{k\geq n} a_k = \sup_{n\in\mathbb{N}} \inf_{k\geq n} a_k. \tag{2.1.1}$$

Similarly the sequence (c_n) where $c_n = \sup_{k\geq n} a_k$ is monotonic decreasing and bounded below. So it also converges to a limit which we call the *limit superior* or *limsup* for short. We denote $\limsup_{n\to\infty} a_n = \lim_{n\to\infty} b_n$. Then we have

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \sup_{k\geq n} a_k = \inf_{n\in\mathbb{N}} \sup_{k\geq n} a_k. \tag{2.1.2}$$

Clearly we have $\liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n$. In Problem 11, you can investigate some other properties of lim sup and lim inf. It can be shown that the smallest limit of any convergent subsequence of a_n is $\liminf_{n\to\infty}$ and the largest limit is $\limsup_{n\to\infty}$. The next theorem is very useful:

Theorem 2.1.1 *A bounded sequence of real numbers (a_n) converges to a limit if and only if $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$. In this case we have*

$$\lim_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n.$$

Proof. If (a_n) converges to a limit, then all of its subsequences also converge to the same limit and it follows that $\lim_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$. Conversely suppose that we don't know that (a_n) converges but we do know that $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$. Then for all $n \in \mathbb{N}$,

$$0 \leq a_n - \inf_{k\geq n} a_k \leq \sup_{k\geq n} a_k - \inf_{k\geq n} a_k.$$

But $\lim_{n\to\infty} \left(\sup_{k\geq n} a_k - \inf_{k\geq n} a_k\right) = \limsup_{n\to\infty} a_n - \liminf_{n\to\infty} a_n = 0$ and so $\lim_{n\to\infty} (a_n - \inf_{k\geq n} a_k) = 0$ by the sandwich rule. But since

$$a_n = \left(a_n - \inf_{k\geq n} a_k\right) + \inf_{k\geq n} a_k,$$

and $\lim_{n\to\infty} \inf_{k\geq n} a_k = \liminf_{n\to\infty} a_n$, we can use the algebra of limits to deduce that (a_n) converges to the common value of $\liminf_{n\to\infty} a_n$ and $\limsup_{n\to\infty} a_n$. □

2.2 Measurable Functions - Basic Concepts

We begin with some motivation from probability. Let $(\Omega, \mathcal{F}, P)$ be a probability space. When we first study probability, we learn that random variables should be considered as mappings from Ω to $\mathbb{R}$. But is this enough for a rigorous mathematical theory? In practise we are interesting in calculating probabilities such as $\text{Prob}(X > a)$ where $a \in \mathbb{R}$. What does this mean in terms of the measure P? We must have

$$\begin{aligned}\text{Prob}(X > a) &= P(\{\omega \in \Omega; X(\omega) \in (a, \infty)\}) \\ & P(X^{-1}((a, \infty))).\end{aligned}$$

Now $X^{-1}((a, \infty)) \subseteq \Omega$, however P only makes sense when applied to sets in $\mathcal{F}$. So we conclude that $\text{Prob}(X > a)$ only makes sense if we impose an additional condition on the mapping X, namely that $X^{-1}((a, \infty)) \in \mathcal{F}$ for all $a \in \mathbb{R}$. This property is precisely what we mean by *measurability.*

In fact let (S, Σ) be an arbitrary measurable space. A mapping $f : S \to \mathbb{R}$ is said to be *measurable* if $f^{-1}((a, \infty)) \in \Sigma$ for all $a \in \mathbb{R}$. So in particular, we should define a *random variable* on a probability space to be a measurable mapping from Ω to $\mathbb{R}$.

Theorem 2.2.1 *Let $f : S \to \mathbb{R}$ be a mapping. The following are equivalent:*

(i) $f^{-1}((a, \infty)) \in \Sigma$,

(ii) $f^{-1}([a, \infty)) \in \Sigma$,

(iii) $f^{-1}((-\infty, a)) \in \Sigma$,

(iv) $f^{-1}((-\infty, a]) \in \Sigma$.

Proof. (i) $\Leftrightarrow$ (iv) as $f^{-1}(A)^c = f^{-1}(A^c)$ and Σ is closed under taking complements.

(ii) $\Leftrightarrow$ (iii) is proved similarly.

(i) $\Rightarrow$ (ii) uses $[a, \infty) = \bigcap_{n=1}^{\infty}(a - 1/n, \infty)$ and so

$$f^{-1}([a, \infty)) = \bigcap_{n=1}^{\infty} f^{-1}((a - 1/n, \infty))$$

and the result follows since Σ is closed under countable intersections.

(ii) $\Rightarrow$ (i) uses

$$f^{-1}((a, \infty)) = \bigcup_{n=1}^{\infty} f^{-1}([a + 1/n, \infty))$$

and the fact that Σ is closed under countable unions. $\square$

It follows that f is measurable if any of (i) to (iv) in Theorem 2.2.1 is established for all $a \in \mathbb{R}$. In Problem 13 you can show that f is measurable if and only if $f^{-1}((a, b)) \in \Sigma$ for all $-\infty \leq a < b \leq \infty$.

A set O in $\mathbb{R}$ is *open* if for every $x \in O$ there is an open interval I containing x for which $I \subseteq O$.

Proposition 2.2.1 *Every open set O in $\mathbb{R}$ is a countable union of disjoint open intervals.*

Proof. For $x \in O$, let I_x be the largest open interval containing x for which $I_x \subseteq O$. If $x, y \in O$ and $x \neq y$ then either I_x and I_y are disjoint or identical, for if they have a non-empty intersection their union is an open interval containing both x and y and that leads to a contradiction unless they coincide. Clearly $O = \bigcup_{x \in O} I_x$. We now select a rational number $r(x)$ in every interval I_x and rewrite O as the countable disjoint union over intervals I_x labelled by distinct rationals $r(x)$. $\square$

It follows immediately from Proposition 2.2.1 that if O is an open set in $\mathbb{R}$ then $O \in \mathcal{B}(\mathbb{R})$.

Theorem 2.2.2 *The mapping $f : S \to \mathbb{R}$ is measurable if and only if $f^{-1}(O) \in \Sigma$ for all open sets O in $\mathbb{R}$.*

Proof. Suppose that $f^{-1}(O) \in \Sigma$ for all open sets O in $\mathbb{R}$. Then in particular $f^{-1}((a, \infty)) \in \Sigma$ for all $a \in \mathbb{R}$ and so f is measurable. Conversely assume that O is open in $\mathbb{R}$ and use Proposition 2.2.1 to write $O = \bigcup_{n=1}^{\infty}(a_n, b_n)$. Then

$$f^{-1}(O) = \bigcup_{n=1}^{\infty} f^{-1}((a_n, b_n)).$$

If f is measurable, then $f^{-1}((a_n, b_n)) \in \Sigma$ for all $n \in \mathbb{N}$ by Problem 13, and the result follows since Σ is closed under countable unions. $\square$

We now present a stronger result than Theorem 2.2.2.

Theorem 2.2.3 *The mapping $f : S \to \mathbb{R}$ is measurable if and only if $f^{-1}(A) \in \Sigma$ for all $A \in \mathcal{B}(\mathbb{R})$.*

Proof. Suppose that f is measurable and let $\mathcal{A} = \{E \subseteq \mathbb{R}; f^{-1}(E) \in \Sigma\}$. We first show that $\mathcal{A}$ is a σ-algebra.

S(i). $\mathbb{R} \in \mathcal{A}$ as $S = f^{-1}(\mathbb{R})$.

S(ii). If $E \in \mathcal{A}$ then $E^c \in \mathcal{A}$ since $f^{-1}(E^c) = f^{-1}(E)^c \in \Sigma$.

S(iii). If (A_n) is a sequence of sets in $\mathcal{A}$ then $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$ since $f^{-1}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \bigcup_{n \in \mathbb{N}} f^{-1}(A_n) \in \Sigma$.

By Problem 13, $f^{-1}((a, b)) \in \Sigma$ for all $-\infty \leq a < b \leq \infty$, and so $\mathcal{A}$ is a σ-algebra of subsets of $\mathbb{R}$ that contains all the open intervals. But by definition, $\mathcal{B}(\mathbb{R})$ is the smallest such σ-algebra. It follows that $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{A}$ and so $f^{-1}(A) \in \Sigma$ for all $A \in \mathcal{B}(\mathbb{R})$. The converse is easy. $\square$

Theorem 2.2.3 leads to the following important extension of the idea of a measurable function: Let (S_1, Σ_1) and (S_2, Σ_2) be measurable spaces. The mapping $f : S_1 \to S_2$ is *measurable* if $f^{-1}(A) \in \Sigma_1$ for all $A \in \Sigma_2$.

Let (S, Σ, m) be a measure space and $f : S \to \mathbb{R}$ be a measurable function. It is easy to see that the mapping $m_f = m \circ f^{-1}$ is a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Indeed $m_f(\emptyset) = 0$ is obvious and if (A_n) is a sequence of disjoint sets in $\mathcal{B}(\mathbb{R})$ we have

$$\begin{aligned} m_f\left(\bigcup_{n=1}^{\infty} A_n\right) &= m\left(f^{-1}\left(\bigcup_{n=1}^{\infty} A_n\right)\right) \\ &= m\left(\bigcup_{n=1}^{\infty} f^{-1}(A_n)\right) \\ &= \sum_{n=1}^{\infty} m(f^{-1}(A_n)) = \sum_{n=1}^{\infty} m_f(A_n), \end{aligned}$$

where we use the fact that for $m \neq n$, $f^{-1}(A_n) \cap f^{-1}(A_m) = f^{-1}(A_n \cap A_m) = f^{-1}(\emptyset) = \emptyset$.

The measure m_f is called the *pushforward* of m by f. In the case of a probability space $(\Omega, \mathcal{F}, P)$ and a random variable $X : \Omega \to \mathbb{R}$, the pushforward is usually denoted p_X. It is a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (you should check that it has total mass 1) and is called the *probability law* or *probability distribution* of the random variable X.

2.3 Examples of Measurable Functions

We first consider the case where $S = \mathbb{R}$ (equipped with its Borel σ algebra) and look for classes of measurable functions. In fact we will prove that

$$\{\text{continuous functions on } \mathbb{R}\} \subseteq \{\text{measurable functions on } \mathbb{R}\}.$$

First we present a result that is well-known (in the wider context of continuous functions on metric spaces) to those who have taken MAS331.

Proposition 2.3.1 *A mapping $f : \mathbb{R} \to \mathbb{R}$ is continuous if and only if $f^{-1}(O)$ is open for every open set O in $\mathbb{R}$.*

Proof. First suppose that f is continuous. Choose an open set O and let $a \in f^{-1}(O)$ so that $f(a) \in O$. Then there exists $\epsilon > 0$ so that $(f(a) - \epsilon, f(a) + \epsilon) \subseteq O$. By definition of continuity of f, for such an ϵ there exists $\delta > 0$ so that $x \in (a - \delta, a + \delta) \Rightarrow f(x) \in (f(a) - \epsilon, f(a) + \epsilon)$. But this tells us that $(a - \delta, a + \delta) \subseteq f^{-1}((f(a) - \epsilon, f(a) + \epsilon)) \subseteq f^{-1}(O)$. Since a is arbitrary we conclude that $f^{-1}(O)$ is open. Conversely suppose that $f^{-1}(O)$ is open for every open set O in $\mathbb{R}$. Choose $a \in \mathbb{R}$ and let $\epsilon > 0$. Then since $(f(a) - \epsilon, f(a) + \epsilon)$ is open so is $f^{-1}((f(a) - \epsilon, f(a) + \epsilon))$. Since $a \in f^{-1}((f(a) - \epsilon, f(a) + \epsilon))$ there exists $\delta > 0$ so that $(a - \delta, a + \delta) \subseteq f^{-1}((f(a) - \epsilon, f(a) + \epsilon))$. From here you can see that whenever $|x - a| < \delta$ we must have $|f(x) - f(a)| < \epsilon$. But then f is continuous at a and the result follows. $\square$

Corollary 2.3.1 *Every continuous function on $\mathbb{R}$ is measurable.*

Proof. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and O be an arbitrary open set in $\mathbb{R}$. Then by Proposition 2.3.1, $f^{-1}(O)$ is an open set in $\mathbb{R}$. Then $f^{-1}(O)$ is in $\mathcal{B}(\mathbb{R})$ by the remark after Proposition 2.2.1. Hence f is measurable by Theorem 2.2.2. $\square$

There are many discontinuous functions on $\mathbb{R}$ that are also measurable. Lets look at an important class of examples in a wider context. Let (S, Σ)

be a general measurable space. Fix $A \in \Sigma$ and define the *indicator function* $\mathbf{1}_A : S \to \mathbb{R}$ by

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

To see that it is measurable its enough to check that

$$\begin{aligned} \mathbf{1}_A^{-1}((c,\infty)) &= \emptyset \in \Sigma & \text{if } c > 1 \\ \mathbf{1}_A^{-1}((c,\infty)) &= A \in \Sigma & \text{if } 0 < c \leq 1 \\ \mathbf{1}_A^{-1}((c,\infty)) &= S \in \Sigma & \text{if } c \leq 0 \end{aligned}$$

If $(S,\Sigma) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ or indeed if S is any metric space, then $\mathbf{1}_A$ is clearly a measurable but discontinuous function.

A particularly interesting example is obtained by taking $(S,\Sigma) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and $A = \mathbb{Q}$. Then $\mathbf{1}_A$ is called *Dirichlet's jump function*. We have already seen that $\mathbb{Q}$ is measurable (it is a countable union of points). As there is a rational number between any pair of irrationals and an irrational number between any pair of rationals, we see that in this case $\mathbf{1}_A$ is measurable, but discontinuous at every point of $\mathbb{R}$.

2.3.1 Algebra of Measurable Functions

One of our goals in this (and the next) section is to show that, sums, products, limits etc of measurable functions are themselves measurable. Throughout this section, (S,Σ) is a measurable space.

Let f and g be functions from S to $\mathbb{R}$ and define for all $x \in S$,

$$(f \vee g)(x) = \max\{f(x), g(x)\} \quad , \quad (f \wedge g)(x) = \min\{f(x), g(x)\}.$$

Proposition 2.3.2 *If f and g are measurable then so are $f \vee g$ and $f \wedge g$.*

Proof. This follows immediately from the facts that for all $c \in \mathbb{R}$,

$$(f \vee g)^{-1}((c,\infty)) = f^{-1}((c,\infty)) \cup g^{-1}((c,\infty))$$

$$\text{and} \quad (f \wedge g)^{-1}((c,\infty)) = f^{-1}((c,\infty)) \cap g^{-1}((c,\infty)) \quad \square$$

Let $-f$ be the function $(-f)(x) = -f(x)$ for all $x \in S$. If f is measurable it is easily checked that $-f$ also is (take $k = -1$ in Problem 14(b).)

Let $\mathbf{0}$ denote the zero function that maps every element of S to zero, i.e. $\mathbf{0} = \mathbf{1}_\emptyset$. Then $\mathbf{0}$ is measurable since it is the indicator factor of a measurable set (or use Problem 12.)

Define $f_+ = f \vee \mathbf{0}$ and $f_- = -f \vee \mathbf{0}$. So that

$$f_+ = \begin{cases} f(x) & \text{if } f(x) \geq 0 \\ 0 & \text{if } f(x) < 0 \end{cases}, f_- = \begin{cases} -f(x) & \text{if } f(x) \leq 0 \\ 0 & \text{if } f(x) > 0 \end{cases}$$

Corollary 2.3.2 *If f is measurable then so are f_+ and f_-.*

Now define the set $\{f > g\} = \{x \in S; f(x) > g(x)\}$.

Proposition 2.3.3 *If f and g are measurable then $\{f > g\} \in \Sigma$.*

Proof Let $\{r_n, n \in \mathbb{N}\}$ be an enumeration of the rational numbers. Then

$$\begin{aligned} \{f > g\} &= \bigcup_{n \in \mathbb{N}} \{f > r_n > g\} \\ &= \bigcup_{n \in \mathbb{N}} \{f > r_n\} \cap \{g < r_n\} \\ &= \bigcup_{n \in \mathbb{N}} f^{-1}((r_n, \infty)) \cap g^{-1}((-\infty, r_n)) \in \Sigma \qquad \square \end{aligned}$$

Theorem 2.3.1 *If f and g are measurable then so is $f + g$.*

Proof. By Problem 14, we see that $a - g$ is measurable for all $a \in \mathbb{R}$. Now

$$(f + g)^{-1}((a, \infty)) = \{f + g > a\} = \{f > a - g\} \in \Sigma,$$

by Proposition 2.3.3 and this establishes the result. $\square$

You can use induction to show that if $f_1, f_2, \ldots, f_n$ are measurable and $c_1, c_2, \ldots, c_n \in \mathbb{R}$ then f is also measurable where $f = c_1 f_1 + c_2 f_2 + \cdots + c_n f_n$. So the set of measurable functions from S to $\mathbb{R}$ forms a real vector space. Of particular interest are the *simple functions* which take the form $f = \sum_{i=1}^n c_i \mathbf{1}_{A_i}$ where $A_i \in \Sigma$ $(1 \leq i \leq n)$. We will learn more about this highly useful class of functions in the next section and in Chapter 3 we will see that they play an important role in integration theory.

Theorem 2.3.2 *If $f : S \to \mathbb{R}$ is measurable and $G : \mathbb{R} \to \mathbb{R}$ is continuous then $G \circ f$ is measurable from S to $\mathbb{R}$.*

Proof. For all $a \in \mathbb{R}$ let $O_a = G^{-1}((a, \infty))$. Then since G is continuous, O_a is an open set in $\mathbb{R}$. Then since for any subset A of S, $(G \circ f)^{-1}(A) = f^{-1}(G^{-1}(A))$, we have

$$(G \circ f)^{-1}((a, \infty)) = f^{-1}(G^{-1}((a, \infty))) = f^{-1}(O_a) \in \Sigma,$$

by Theorem 2.2.2. The result follows. $\square$

Theorem 2.3.3 *If f and g are measurable then so is fg.*

Proof. Apply Theorem 2.3.2 with $G(x) = x^2$ to deduce that h^2 is measurable whenever h is. But

$$fg = \frac{1}{4}[(f+g)^2 - (f-g)^2]$$

and the result follows by using Theorem 2.3.1. $\square$

2.3.2 Limits of Measurable Functions

Let (f_n) be a bounded sequence of functions from S to $\mathbb{R}$ so that $\sup_{n\in\mathbb{N}} \sup_{x\in S} |f_n(x)| < \infty$. Define $\inf_{n\in\mathbb{N}} f_n$ and $\sup_{n\in\mathbb{N}} f_n$ by

$$\left(\inf_{n\in\mathbb{N}} f_n\right)(x) = \inf_{n\in\mathbb{N}} f_n(x) \quad \text{and} \quad \left(\sup_{n\in\mathbb{N}} f_n\right)(x) = \sup_{n\in\mathbb{N}} f_n(x)$$

for all $x \in S$.[1]

Proposition 2.3.4 *If f_n is measurable for all $n \in \mathbb{N}$ then $\inf_{n\in\mathbb{N}} f_n$ and $\sup_{n\in\mathbb{N}} f_n$ are both measurable.*

Proof. For all $c \in \mathbb{R}$,

$$\left\{\inf_{n\in\mathbb{N}} f_n > c\right\} = \bigcap_{n\in\mathbb{N}} \{f_n > c\} \in \Sigma.$$

$$\left\{\sup_{n\in\mathbb{N}} f_n > c\right\} = \bigcup_{n\in\mathbb{N}} \{f_n > c\} \in \Sigma. \qquad \square$$

Define $\liminf_{n\to\infty} f_n$ and $\limsup_{n\to\infty} f_n$ by

$$\left(\liminf_{n\to\infty} f_n\right)(x) = \liminf_{n\to\infty} f_n(x) \quad \text{and} \quad \left(\limsup_{n\to\infty} f_n\right)(x) = \limsup_{n\to\infty} f_n(x)$$

for all $x \in S$

Theorem 2.3.4 *If f_n is measurable for all $n \in \mathbb{N}$ then $\liminf_{n\to\infty} f_n$ and $\limsup_{n\to\infty} f_n$ are both measurable.*

[1] We can drop the boundedness requirement if we work with functions taking values in $[-\infty, \infty]$.

Proof. By Proposition 2.3.4, $\inf_{k\geq n} f_k$ and $\sup_{k\geq n} f_k$ are measurable for each $n \in \mathbb{N}$. Then by Proposition 2.3.4 again, $\liminf_{n\to\infty} f_n = \sup_{n\in\mathbb{N}} \inf_{k\geq n} f_k$ and $\limsup_{n\to\infty} f_n = \inf_{n\in\mathbb{N}} \sup_{k\geq n} f_k$ are measurable. $\square$

We say that the sequence (f_n) converges *pointwise* to f as $n \to \infty$ if $\lim_{n\to\infty} f_n(x) = f(x)$ for all $x \in S$.

Theorem 2.3.5 *If If f_n is measurable for all $n \in \mathbb{N}$ and (f_n) converges pointwise to f as $n \to \infty$, then f is measurable.*

Proof. By Theorem 2.1.1 $f(x) = \liminf_{n\to\infty} f_n(x)$ for all $x \in S$ and so f is measurable by Theorem 2.3.4. $\square$

2.4 Simple Functions

Recall the definition of indicator functions $\mathbf{1}_A$ where $A \in \Sigma$. A mapping $f : S \to \mathbb{R}$ is said to be *simple* if it takes the form

$$f = \sum_{i=1}^{n} c_i \mathbf{1}_{A_i} \tag{2.4.3}$$

where $c_1, c_2, \ldots, c_n \in \mathbb{R}$ and $A_1, A_2, \ldots, A_n \in \Sigma$ with $\bigcup_{i=1}^{n} A_i = S$ and $A_i \cap A_j = \emptyset$ when $i \neq j$. In other words a simple function is a (finite) linear combination of indicator functions of non-overlapping sets. It follows from Theorem 2.3.1 that every simple function is measurable. It is straightforward to prove that sums and scalar multiples of simple functions are themselves simple, so the set of all simple functions form a vector space.

We now prove a key result that shows that simple functions are powerful tools for approximating measurable functions. Recall that a mapping $f : S \to \mathbb{R}$ is *non-negative* if $f(x) \geq 0$ for all $x \in S$, which we write for short as $f \geq 0$. We write $f \leq g$ when $g - f \geq 0$. It is easy to see that a simple function of the form (2.4.3) is non-negative if and only if $c_i \geq 0$ $(1 \leq i \leq n)$.

Theorem 2.4.1 *Let $f : S \to \mathbb{R}$ be measurable and non-negative. Then there exists a sequence (s_n) of non-negative simple functions on S with $s_n \leq s_{n+1} \leq f$ for all $n \in \mathbb{N}$ so that (s_n) converges pointwise to f as $n \to \infty$. If f is bounded then the convergence is uniform.*

Proof. We split this into three parts.

Step 1 Construction of (s_n).

Divide the interval $[0, n)$ into $n2^n$ subintervals $\{I_j, 1 \leq j \leq n2^n\}$, each of length $\frac{1}{2^n}$ by taking $I_j = \left[\frac{j-1}{2^n}, \frac{j}{2^n}\right)$. Let $E_j = f^{-1}(I_j)$ and $F_n = f^{-1}([n, \infty))$. Then $S = \bigcup_{j=1}^{n2^n} E_j \cup F_n$. We define for all $x \in S$

$$s_n(x) = \sum_{j=1}^{n2^n} \left(\frac{j-1}{2^n}\right) \mathbf{1}_{E_j}(x) + n\mathbf{1}_{F_n}(x).$$

Step 2 Properties of (s_n).

For $x \in E_j, s_n(x) = \frac{j-1}{2^n}$ and $\frac{j-1}{2^n} \leq f(x) < \frac{j}{2^n}$ and so $s_n(x) \leq f(x)$. For $x \in F_n, s_n(x) = n$ and $f(x) \geq n$. So we conclude that $s_n \leq f$ for all $n \in \mathbb{N}$.

To show that $s_n \leq s_{n+1}$, fix an arbitrary j and consider $I_j = \left[\frac{j-1}{2^n}, \frac{j}{2^n}\right)$. For convenience, we write I_j as I and we observe that $I = I_1 \cup I_2$ where $I_1 = \left[\frac{2j-2}{2^{n+1}}, \frac{2j-1}{2^{n+1}}\right)$ and $I_2 = \left[\frac{2j-1}{2^{n+1}}, \frac{2j}{2^{n+1}}\right)$. Let $E = f^{-1}(I), E_1 = f^{-1}(I_1)$ and $E_2 = f^{-1}(I_2)$. Then $s_n(x) = \frac{j-1}{2^n}$ for all $x \in E, s_{n+1}(x) = \frac{j-1}{2^n}$ for all $x \in E_1$, and $s_{n+1}(x) = \frac{2j-1}{2^{n+1}}$ for all $x \in E_2$. It follows that $s_n \leq s_{n+1}$ for all $x \in E$. A similar (easier) argument can be used on F_n.

Step 3 Convergence of (s_n).

Fix any $x \in S$. Since $f(x) \in \mathbb{R}$ there exists $n_0 \in \mathbb{N}$ so that $f(x) \leq n_0$. Then for each $n > n_0, f(x) \in I_j$ for some $1 \leq j \leq n2^n$, i.e. $\frac{j-1}{2^n} \leq f(x) < \frac{j}{2^n}$. But $s_n(x) = \frac{j-1}{2^n}$ and so $|f(x) - s_n(x)| < \frac{1}{2^n}$ and the result follows. If f is bounded we can find $n_0 \in \mathbb{N}$ so that $f(x) \leq n_0$ for all $x \in \mathbb{R}$. Then the argument just given yields $|f(x) - s_n(x)| < \frac{1}{2^n}$ for all $x \in \mathbb{R}$ from which we can deduce the uniformity of the convergence. $\square$