Say that a family $\mathcal{M}$ of subsets of a non-empty set Ω is a *monotone class* if the following holds: if $A_n \uparrow A$ or $A_n \downarrow A$ and each $A_n \in \mathcal{M}$ then $A \in \mathcal{M}$.

Here $A_n \uparrow A$ means that $A_1 \subset A_2 \subset A_3 \subset \cdots$ and $A = \cup_n A_n$. And $A_n \downarrow A$ means that $A_1 \supset A_2 \supset A_3 \supset \cdots$ and $A = \cap_n A_n$.

For example if Ω is the set of real numbers, then the collection of all intervals (empty, degenerate, open, closed, semiopen) is a monotone class. Given any class $\mathcal{C}$ of subsets of Ω, one can show that there is a smallest monotone class $\mathcal{M}(\mathcal{C})$ of subsets of Ω that includes the given class. In fact consider all possible collections which include the given family of sets and form a monotone family. For example, the class of all subsets of Ω is one such.

Now consider those sets which are in all the monotone classes we considered above. This collection is a monotone class and is the smallest monotone class that includes the given collection of sets. $\mathcal{M}(\mathcal{C})$ is called the monotone class generated by $\mathcal{C}$.

Recall $\sigma(\mathcal{C})$ is the smallest sigma field which includes a given class of sets $\mathcal{C}$. $\sigma(\mathcal{C})$ is called the sigma field generated by $\mathcal{C}$.

Theorem : Monotone Class Theorem, MCT
If $\mathcal{F}$ is a field of subsets of Ω then $\sigma(\mathcal{F}) = \mathcal{M}(\mathcal{F})$.

proof : A σ-field is a monotone class — in particular $\sigma(\mathcal{F})$ is a monotone class. Since $\mathcal{M}(\mathcal{F})$ is the smallest monotone class that includes $\mathcal{F}$ we conclude that $\mathcal{M}(\mathcal{F}) \subset \sigma(\mathcal{F})$ It remains to show that $\mathcal{M}(\mathcal{F}) \supset \sigma(\mathcal{F})$.

We now show that $\mathcal{M}(\mathcal{F})$ is a field. What does this achieve? Since a field which is a monotone class is a σ-field, this shows that $\mathcal{M}(\mathcal{F})$ is indeed a σ-field. It then follows that $\mathcal{M}(\mathcal{F})$ is a σ-field that includes the class $\mathcal{F}$, so must include the smallest such namely, $\sigma(\mathcal{F})$.

First, we show that whenever $A \in \mathcal{M}(\mathcal{F})$ then so is A^c. For this, consider the class $\mathcal{M}_0 = \{A \in \mathcal{M}(\mathcal{F}) : A^c \in \mathcal{M}(\mathcal{F})\}$. If $A \in \mathcal{F}$, then $\mathcal{F}$ being a field we have $A^c \in \mathcal{F}$ which is included in $\mathcal{M}(\mathcal{F})$, it follows that $\mathcal{F} \subset \mathcal{M}_0$. If each A_n is in $\mathcal{M}_0$ and $A_n \uparrow A$, then firstly A_n as well as A_n^c are in $\mathcal{M}(\mathcal{F})$. Secondly, $A_n \uparrow A$ as well as $A_n^c \downarrow A^c$ so that both A and A^c are in $\mathcal{M}(\mathcal{F})$. Thus $A \in \mathcal{M}_0$. Similarly, if each A_n is in $\mathcal{M}_0$ and $A_n \downarrow A$ then $A \in \mathcal{M}_0$. In other words $\mathcal{M}_0$ is a monotone class that includes $\mathcal{F}$. But $\mathcal{M}(\mathcal{F})$ is the smallest such. This means that every set in $\mathcal{M}(\mathcal{F})$ must be in $\mathcal{M}_0$. Thus, whenever $A \in \mathcal{M}(\mathcal{F})$ so is A^c.

Second, we show that if A and B are in $\mathcal{M}(\mathcal{F})$ then so is their intersection $A \cap B$. To start with, observe that if both the sets A and B are in $\mathcal{F}$ then their intersection is already in $\mathcal{F}$ and hence in $\mathcal{M}(\mathcal{F})$. Fix $A \in \mathcal{F}$. Consider $\mathcal{M}_0 = \{B \in \mathcal{M}(\mathcal{F}) : A \cap B \in \mathcal{M}(\mathcal{F})\}$. Beacuse of the earlier sentence, this family includes $\mathcal{F}$. As in the earlier paragraph, $\mathcal{M}_0$ is a monotone class and hence must be $\mathcal{M}(\mathcal{F})$. In other words, for every $B \in \mathcal{M}(\mathcal{F})$ we have $A \cap B \in \mathcal{M}(\mathcal{F})$. Now fix any $B \in \mathcal{M}(\mathcal{F})$ and consider $\mathcal{M}_0 = \{A \in \mathcal{M}(\mathcal{F}) : A \cap B \in \mathcal{M}(\mathcal{F})\}$. From what we proved just now, this family includes $\mathcal{F}$. This is again a monotone class and hence must be $\mathcal{M}(\mathcal{F})$. Thus intersection of two sets in $\mathcal{M}(\mathcal{F})$ is again in $\mathcal{M}(\mathcal{F})$.

This completes the proof.

Measurable functions:

Let R be the real line and let $\mathcal{B}$ be the sigma field generated by the open subsets of R; equivalently the sigma field generated by the closed subsets of R. $\mathcal{B}$ is called the *Borel σ field of R*.

Theorem

$\mathcal{B} = \sigma(\mathcal{E}_i)$ if $\mathcal{E}_i$ is any of the following:
(1) $\mathcal{E}_1 =$ the open intervals $\{(a, b) : a < b, a, b \in R\}$.
(2) $\mathcal{E}_2 =$ the closed intervals $\{[a, b] : a < b, a, b \in R\}$.
(3) The half-open intervals $\mathcal{E}_3 =\{(a, b] : a < b, a, b \in R\}$, or $\mathcal{E}_4 = \{[a, b) : a < b, a, b \in R\}$.
(4) Open rays $\mathcal{E}_5 = \{(a, \infty), a \in R\}$, or $\mathcal{E}_6= \{(-\infty, a), a \in R\}$.
(5) Closed rays $\mathcal{E}_7=\{[a, \infty), a \in R\}$, or $\mathcal{E}_8=\{(-\infty, a], a \in R\}$.

Suppose now that Ω is a set and $\mathcal{A}$ is a sigma field of subsets of Ω.
Let L be the collection of all real (Borel) measurable functions on Ω. Recall f is measurable means that for every $B \in \mathcal{B}$, $f^{-1}(B) \in \mathcal{A}$.

Theorem:

If $\sigma(\mathcal{E}) = \mathcal{B}$, then $f : \Omega \mapsto R$ is measurable if, and only if, $f^{-1}(E) \in \mathcal{A}$ for each $E \in \mathcal{E}$.

So $f : \Omega \mapsto R$ is measurable if, and only if, for every $a \in R$, the set $\{\omega : f(\omega) \le a\} \in \mathcal{A}$.
Or, $f : \Omega \mapsto R$ is measurable if, and only if, for every $a \in R$, the set $\{\omega : f(\omega) < a\} \in \mathcal{A}$.

Among measurable functions let us identify the simplest ones : A measurable function X is *simple measurable function* if it takes finitely many real values. We denote by I_A the indicator function of the set A, namely the function which is 1 for points of A and zero for points not in A.

Theorem : Structure of real measurable functions;

(i) The space L is a real vector space, that is, if X and Y are in L, then so is $\alpha X + \beta Y$ for any real numbers α and β. L is an algebra, that is, besides being a vector space it is closed under multiplication. L is a lattice too, that is, the (pointwise) maximum and minimum of two measurable functions is again a measurable function. Constant functions are in L. Thus in particular, for any measurable function X, both $X^+ = \max(X, 0)$ and $X^- = \max(-X, 0)$ are measurable functions. Clearly, $X = X^+ - X^-$.
(ii) If $X_n \in L$ and $X_n \uparrow X$ or $X_n \downarrow X$ pointwise and X is real valued, then $X \in L$. In particular pointwise $\limsup$ and $\liminf$ of X_n are in L provided they are real valued.

(iii) $I_A \in L$ iff $A \in \mathcal{A}$. The collection of simple measurable functions is also a vector space, algebra and lattice – as in (i) above.

(iv) Given any bounded measurable function X there is a sequence (s_n) of simple measurable functions such that $s_n \uparrow X$ uniformly. Given any non-negative measurable function X, there is a sequence (s_n) of simple non-negative measurable functions such that $s_n \uparrow X$ point wise.

(v) Suppose that $\mathcal{F}$ is a field of sets such that $\mathcal{A} = \sigma(\mathcal{F})$. Suppose that Φ is a vector space of real valued functions on Ω closed under monotone limits — which means that if each $X_n \in \Phi$ and $X_n \uparrow X$ or $X_n \downarrow X$ point wise and X is real valued, then $X \in \Phi$. Suppose that for each $A \in \mathcal{F}$, we have $I_A \in \Phi$. Then $L \subset \Phi$.

(vi) Suppose that $\mathcal{F}$ is a field of sets such that $\mathcal{A} = \sigma(\mathcal{F})$. Suppose that Φ is a vector space of real valued functions on Ω closed under bounded point wise convergence — which means that if each $X_n \in \Phi$ and $X_n \to X$ pointwise and X_n are uniformly bounded, then $X \in \Phi$. Suppose that for each $A \in \mathcal{F}$, we have $I_A \in \Phi$. Then every bounded measurable function is in Φ.

It is rather awkward that we needed to assume in (ii) above that the lim sup etc are real valued. Actually this is unnecessary. But we have to admit extended real valued functions to make things precise. We shall do this soon. The last two parts of the theorem above help us in dealing with the collection of all measurable functions just as the MCT helps us in dealing with all sets in the σ-field.

Proof: To prove X is a measurable function it suffices to show that for any number a, the set $X^{-1}(-\infty, a)$ is in our σ-field, or, for any number a, the set $X^{-1}(-\infty, a]$ is in our σ-field.

(i) For any number a, $(X + Y < a) = \cup_r \{(X < r) \cap (Y < a - r)\}$ where the union runs over the countable set of rationals r. Thus if X and Y are measurable functions then so is $X + Y$. For $\alpha > 0$, $(\alpha X < a) = (X < a/\alpha)$ and for $\alpha < 0$, $(\alpha X < a) = (X > a/\alpha)$ to conclude that αX is a measurable functionif X is so. Since $(X^2 \le a)$ is empty set if $a < 0$ where as it equals $(-\sqrt{a} \le X \le +\sqrt{a})$ for $a > 0$ showing that X^2 is a measurable function if X is so. Thus in particular combining these it follows that if X and Y are measurable functions then so is $XY = [(X+Y)^2 - X^2 - Y^2]/2$. Observe that $(X \wedge Y < a) = (X < a) \cup (Y < a)$ and $(X \vee Y < a) = (X < a) \cap (Y < a)$ so that $X \wedge Y$ and $X \vee Y$ are measurable functions if X and Y are so. Since constant functions are visibly measurable functions this proves (i).

(ii) If $X_n \uparrow X$ then $(X \le a) = \cap (X_n \le a)$. If $X_n \downarrow X$ then $(X < a) = \cup (X_n < a)$ to prove the first statement. Let X_n be measurable functions. Fix n. For each $m > n$, $\max\{X_n, X_{n+1}, \cdots, X_m\}$ is a measurable function by (i) and increases to, say, Y_n as $m \uparrow \infty$. By the first statement each Y_n is a measurable function. But then $Y_n \downarrow \limsup X_n$ so that this last one is also a measurable function. Similar argument holds for liminf.

(iii) Since $(I_A < 1) = A^c$ it follows that I_A is a measurable function iff A is in our σ-field. For the second statement, you only need to verify that sums ,products, scalar multiples of simple functions are simple again.

(iv) For convenience assume that $0 \le X \le 1$. Define $s_n(\omega)$ as follows: if $0 \le X(\omega) \le 1/2^n$ then $s_n(\omega) = 0$ and for $1 \le k \le 2^n - 1$ put $s_n(\omega) = k/2^n$ in case $k/2^n < X(\omega) \le (k+1)/2^n$. This sequence does.

Given non-negative X and an integer $n \ge 1$ define s_n as follows: $s_n(\omega) = 0$ in case $0 \le X(\omega) \le 1/2^n$; $s_n(\omega) = n$ in case $X(\omega) > n$. For $1 \le k \le n2^n - 1$, if $k/2^n < X(\omega) \le (k+1)/2^n$ put $s_n(\omega) = k/2^n$. This sequence does.

(v) First observe that the class of sets A such that $I_A \in \Phi$ is a monotone class and includes $\mathcal{F}$ — all this by hypothesis. So by MCT, we conclude that $I_A \in \Phi$ for each $A \in \mathcal{A}$. Since Φ is a vector space, it must include every simple measurable function. By hypothesis and (iv) it must include every nonnegative measurable function too. For any measurable function X, since $X = X^+ - X^-$ we conclude that X must also be in Φ to complete proof.

(vi) The required ideas are already in (v) above.

Some inequalities

If X is a measurable function on $(\Omega, \mathcal{A}, \mu)$, then for $1 \le p < \infty$,

$$\|X\|_p = (\int |X|^p d\mu)^{1/p}.$$

$$L^p = \{X : \Omega \mapsto R, \|X\|_p < \infty\}.$$

Holder's inequality : For nonnegative measurable functions X and Y, and for numbers $p > 1$, $q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$; $\|XY\|_1 \le \|X\|_p \cdot \|Y\|_q$. In particular, if $X \in L^p$ and $Y \in L^q$, then $XY \in L^1$, and in this case the inequality is an equality iff $\alpha|X|^p = \beta|Y|^q$ a.e. for some constants α, β with $\alpha\beta \ne 0$

Cauchy-Schwarz inequality : The above inequality when $p = 2$ and $q = 2$ – that is, for measurable functions X and Y; $\|XY\|_1 \le \|X\|_2 \|Y\|_2$

Minkowski inequality : For measurable functions X and Y, and $1 \le p < \infty$, $\|X + Y\|_p \le \|X\|_p + \|Y\|_p$.

First, let us show that Minkowski's inequality fails when $p < 1$. For $p < 1$, if $a > 0, b > 0, t > 0$ then $t^{p-1} > (a+t)^{p-1}$. Integrating this from 0 to b, we get $b^p/p > (a+b)^p/p - a^p/p$. So, $a^p + b^p > (a+b)^p$. Now, suppose E, F are disjoint subsets of finite measure, $a = (\mu(E))^{1/p}, b = (\mu(F))^{1/p}$, then, $(\int (I_E + I_F)^p)^{1/p} = (\mu(E) + \mu(F))^{1/p} = (a^p + b^p)^{1/p} > a + b = (\mu(E))^{1/p} + (\mu(F))^{1/p} = (\int (I_E)^p)^{1/p} + (\int (I_F)^p)^{1/p}$.

Lemma:
If $a \ge 0, b \ge 0$, and $0 < \lambda < 1$, then,

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda) b,$$

with equality if, and only if $a = b$.
Proof: This follows from the strict convexity of the exponential function $\exp(x)$.

Proof of Holder's inequality

The result is trivial, if $X = 0$ a.e., or if $Y = 0$ a.e. The result is also trivial is $\int X^p = \infty$ or if $\int Y^q = \infty$. Otherwise, we apply the lemma with

$$a = \frac{(X(\omega))^p}{\int X^p}, \;\; b = \frac{(Y(\omega))^q}{\int Y^q}, \;\; \lambda = 1/p$$

to obtain,

$$\frac{X(\omega)Y(\omega)}{\|X\|_p \cdot \|Y\|_q} \leq \frac{X(\omega)^p}{p\int X^p} + \frac{(Y(\omega)^q}{q\int Y^q}$$

,
for every ω. Integration of both sides yields,

$$\frac{\int XY}{\|X\|_p \cdot \|Y\|_q} \leq 1/p + 1/q = 1.$$

Equality holds above, iff,

$$\frac{X(\omega)Y(\omega)}{\|X\|_p \cdot \|Y\|_q} = \frac{X(\omega)^p}{p\int X^p} + \frac{(Y(\omega)^q}{q\int Y^q}$$

,

for a. e. ω. This happens precisely when $a = b$, or

$$(\int Y^q) \cdot X^p = (\int X^p) \cdot Y^q$$

a.e.

Note: If $1 < p < \infty$, the number q such that $1/p + 1/q = 1$ is called the conjugate exponent of p.

Proof of Minkowski's inequality:

For a measurable function X and $1 \leq p < \infty$, let $\|X\|_p = (\int |X|^p)^{1/p}$. Now, Minkowski's inequlity is obvious if $p = 1$ or if $X + Y = 0$ a.e. Otherwise,

$$|X+Y|^p \leq (|X| + |Y|)(|X+Y|^{p-1}).$$

Apply Holder's inequlity.

$$\int |X+Y|^p \leq \|X\|_p \||X+Y|^{p-1}\|_q + \|Y\|_p \||X+Y|^{p-1}\|_q$$

Now, $(p-1)q = p$. So,

$$\|X+Y\|_p = (\int |X+Y|^p)^{1-1/q} \leq \|X\|_p + \|Y\|_q.$$

For a measurable function X, we define,

$$\|X\|_\infty = \inf\{a \geq 0 : \mu(\{\omega : |X(\omega)| > a\}) = 0\}$$

with the convention that $\inf \emptyset = \infty$. Easy to show that $\mu(\{\omega : |X(\omega)| > \|X\|_\infty\} = 0$. $\|X\|_\infty$ is called the *essential supremum of* X.