By definition, a function $f:\Omega \to [-\infty,\infty]$ is a random variable (is "measurable") iff $f^{-1}(B) \in \mathcal{F}$ for every Borel(-Lebesgue) set $B \subseteq [-\infty,\infty]$.

Proposition

Let $f:\Omega \to [-\infty,\infty]$ satisfy the "level set condition" $\forall \lambda \in [-\infty,\infty]$, $f^{-1}([-\infty,\lambda]) \in \mathcal{F}$. Then f is measurable.
Conversely, every measurable function satisfies the level set condition.

Pf

Let $\mathcal{B} = \{B \subseteq [-\infty,\infty] : f^{-1}(B) \in \mathcal{F}\}$. Then

(i) $[-\infty,\infty] \in \mathcal{B}$

(ii) If $B \in \mathcal{B}$ then $f^{-1}([-\infty,\infty]\setminus B) = f^{-1}([-\infty,\infty])\setminus f^{-1}(B)$ $= \Omega \setminus f^{-1}(B) \in \mathcal{F}$, so B is closed under complementation.

(iii) If $B_1, B_2 \in \mathcal{B}$ then $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$ $\in \mathcal{F}$. So $\mathcal{B}$ is closed under (finite) intersections.

(iv) If $(B_n) \subset \mathcal{B}$ then $f^{-1}\left(\bigcup_n B_n\right) = \bigcup_n f^{-1}(B_n) \in \mathcal{F}$. So $\mathcal{B}$ is closed under countable unions.

So $\mathcal{B}$ is a σ-field. From the level set property; $(\lambda, \mu] \in \mathcal{B}$ for $\lambda, \mu \in [-\infty,\infty]$. A short argument shows $\{\lambda\} \in \mathcal{B}$ for $\lambda \in [-\infty,\infty]$ and consequently any kind of interval belongs to $\mathcal{B}$. Now $\mathcal{B}$ is a σ-field and the σ-field generated by the intervals in $[-\infty,\infty]$, the Borel σ-field, must be contained in $\mathcal{B}$. So f is measurable

Lemma

Let $f:\Omega \to [-\infty,\infty]$ be measurable and $g:\mathbb{R}\to\mathbb{R}$ a Borel measurable function. Let $\mathcal{B}$ denote the Borel σ-field on $\mathbb{R}$. The $g\circ f$ is measurable too.

Pf Let $H \in \mathcal{B}$ then $g^{-1}(H) \in \mathcal{B}$ and so $f^{-1}(g^{-1}(H)) \in \mathcal{F}$.
But $f^{-1}(g^{-1}(H)) = (g\circ f)^{-1}(H)$. □

Lemma

If f is measurable and $f^+ = \max\{f,0\}$, $f^- = \max\{-f,0\}$ while $|f| = f^+ + f^-$. Then $f^+, f^-, |f|$ are all measurable.

Pf Set $g^+(x) = \max\{x,0\}$, $g^-(x) = \max\{-x,0\}$, $g(x) = |x|$. Each of these are ~~cts~~ fns(†). So, by the previous lemma, $g^+(f) = f^+$, $g^-(f) = f^-$ and $g(f) = |f|$ are measurable.

Lemma

Suppose that $f:\Omega \to [-\infty,\infty]$ is the pointwise limit of simple random variables, s_n. So $\forall \omega\in\Omega$, $s_n(\omega) \to f(\omega)$. Then f is (a random variable) measurable.

Pf Let $m \in \mathbb{N}$, then, as $s_n(\omega) \to f(\omega)$, should $\omega \in \{f \le \lambda\}$ we have $\omega \in \bigcap_{m=1}^{\infty}\bigcup_{\ell=1}^{\infty}\bigcap_{n\ge\ell}\{s_n \le \lambda + 1/m\}$. This last set belongs to $\mathcal{F}$. On the other hand, if $\omega \in \bigcap_{m=1}^{\infty}\bigcup_{\ell=1}^{\infty}\bigcap_{n\ge\ell}\{s_n \le \lambda + 1/m\}$ then for each $m\in\mathbb{N}$ there is $\ell = \ell(\omega,m)$ such that $\omega \in \bigcap_{n\ge\ell}\{s_n \le \lambda + \frac{1}{m}\}$ and therefore $f(\omega) = \lim_n s_n(\omega) \le \lambda + 1/m$. So $f(\omega) \le \lambda$. Thus

$$\{f \le \lambda\} = \bigcap_{m=1}^{\infty}\bigcup_{\ell=1}^{\infty}\bigcap_{n\ge\ell}\{s_n \le \lambda + \tfrac{1}{m}\} \in \mathcal{F}.$$

So f is measurable.

(†) Hence Borel measurable.

Lemma

Let f, g be random variables. Then $f+g$ is a random variable. Also if $\mu \in \mathbb{R}$, then μf is a random variable.

Pf $f+g = f^+ - f^- + g^+ - g^-$. Each of f^+, f^-, g^+, g^- are random variables and we know that each of these is the limit of a (n increasing) sequence of simple random variables. Any linear combination of these random variables is the limit of the sequence formed by the corresponding linear combination of simple random variables — which is again a sequence of simple random variables. So $f+g$ is a random variable by the previous lemma.

If $\mu = 0$, μf is a simple random variable. If $\mu > 0$ then $\{\mu f \leq \lambda\} = \{f \leq \frac{\lambda}{\mu}\} \in \mathcal{F}$. If $\mu < 0$ note that $\{f \leq \frac{\lambda}{\mu}\} = \{\mu f \geq \lambda\} \in \mathcal{F}$. So $\{\mu f < \lambda\} \in \mathcal{F}$ and $\{f = \frac{\lambda}{\mu}\} = f^{-1}\{\frac{\lambda}{\mu}\} \in \mathcal{F}$. So $\{\mu f \leq \lambda\} = \{\mu f < \lambda\} \cup f^{-1}\{\frac{\lambda}{\mu}\} \in \mathcal{F}$.

Lemma

Let f_1, f_2 be random variables, then so are,

$$\max\{f_1, f_2\} \triangleq f_1 \vee f_2 = \frac{f_1+f_2}{2} + \frac{|f_1-f_2|}{2}$$

$$\min\{f_1, f_2\} \triangleq f_1 \wedge f_2 = \frac{f_1+f_2}{2} - \frac{|f_1-f_2|}{2}$$

Pf Previous lemmas.

Definition Let f be a random variable. We define $\mathbb{E}_{\mathbb{P}}(f)$, the expectation of f with respect to $\mathbb{P}$, to be

$$\mathbb{E}_{\mathbb{P}}(f) \triangleq \int_\Omega f \, d\mathbb{P} \triangleq \int_\Omega f^+ d\mathbb{P} - \int f^- d\mathbb{P}.$$

For any $f = g-h$, $g>0$ & h we define $\int f \, d\mathbb{P} \overset{\Omega}{=} \int g \, d\mathbb{P} - \int h \, d\mathbb{P}$

So long as at least one of the right hand integrals is finite.

<u>Lemma</u>

Suppose that $f:\Omega \to [-\infty, \infty]$ is a random variable and $f = g - h$, where $g \geq 0 \leq h$ are +ve random variables, and, also, $f = u - v$, where $u \geq 0 \leq v$ are +ve random variables too. Then,

$$\int_\Omega g \, d\mathbb{P} - \int_\Omega h \, d\mathbb{P} = \int_\Omega u \, d\mathbb{P} - \int_\Omega v \, d\mathbb{P}.$$

<u>Pf</u> Since $g - h = f = u - v$ then $g + v = u + h$. By our Corollary to the Monotone Convergence Theorem (for positive random variables),

$$\int_\Omega g\,d\mathbb{P} + \int_\Omega v\,d\mathbb{P} = \int_\Omega (g+v)\,d\mathbb{P} = \int_\Omega (u+h)\,d\mathbb{P} = \int_\Omega u\,d\mathbb{P} + \int_\Omega h\,d\mathbb{P},$$

so $\int_\Omega g\,d\mathbb{P} - \int_\Omega h\,d\mathbb{P} = \int_\Omega u\,d\mathbb{P} - \int_\Omega v\,d\mathbb{P}$. □

<u>Corollary</u>

The integral of a random variable with respect to $\mathbb{P}$ is 'well defined'. The integral is a linear functional.

<u>Pf</u>

Write $f = f^+ - f^-$ and $g = g^+ - g^-$ then $f + g = f^+ + g^+ - (f^- + g^-)$ and

$$\int_\Omega (f+g)\,d\mathbb{P} = \int_\Omega (f^+ + g^+)\,d\mathbb{P} - \int_\Omega (f^- + g^-)\,d\mathbb{P} \quad \text{using the definition and the Lemma,}$$

$$= \int_\Omega f^+ d\mathbb{P} + \int_\Omega g^+ d\mathbb{P} - \int_\Omega f^- d\mathbb{P} - \int_\Omega g^- d\mathbb{P} \quad \text{by result for +ve r.v.'s}$$

$$= \int_\Omega f^+ d\mathbb{P} - \int_\Omega f^- d\mathbb{P} + \int_\Omega g^+ d\mathbb{P} - \int_\Omega g^- d\mathbb{P}$$

$$= \int_\Omega f\,d\mathbb{P} + \int_\Omega g\,d\mathbb{P}.$$ The other bit is easier. □

<u>The Monotone Convergence Theorem</u> (For 'general' r.v.'s)

Let (f_n) be a sequence of r.v.'s such that for $f : \Omega \to [-\infty, \infty]$

(i) $f_1 \leq f_2 \leq \ldots \leq f$

(ii) $f_n(\omega) \underset{n\to\infty}{\longrightarrow} f(\omega) \quad \forall \omega \in \Omega$.

(iii) $\int_\Omega f_n \, d\mathbb{P}$ exists for each $n \in \mathbb{N}$, and is finite.

Then f is a r.v. and $\sup_n \int_\Omega f_n \, d\mathbb{P} = \int_\Omega f \, d\mathbb{P}$

<u>Pf</u>

Each f_n satisfies the level set condition, $\{f_n \leq \lambda\} \in \mathcal{F}$, $\forall \lambda \in [-\infty, \infty]$. Let $E_n = \{f_n \leq \lambda\}$. Then $\{f \leq \lambda\} = \bigcap_n E_n \in \mathcal{F}$. So f is a r.v. Consider a sequence, (g_n), of r.v.'s which are non-negative and decreasing and converging pointwise to 0. Then the sequence $(g_1 - g_n)$ is non-negative, increasing, comprises r.v.'s. Suppose also that $\int_\Omega g_1 \, d\mathbb{P} < \infty$. Then so are $\int_\Omega g_n \, d\mathbb{P}$, $n \geq 1$. By our original version of the monotone convergence theorem; since $\sup_n (g_1 - g_n) = g_1$, then $\sup_n \int_\Omega (g_1 - g_n) \, d\mathbb{P} = \int_\Omega g_1 \, d\mathbb{P}$, that is,

$$\int_\Omega g_1 \, d\mathbb{P} - \inf_n \int_\Omega g_n \, d\mathbb{P} = \int_\Omega g_1 \, d\mathbb{P}$$

So $\inf_n \int_\Omega g_n \, d\mathbb{P} = 0$. Applying this to $(f - f_n)$ we get

$$\int_\Omega f \, d\mathbb{P} = \sup_n \int_\Omega f_n \, d\mathbb{P}$$

If $f : \Omega \to [-\infty, \infty]$ is a random variable and $f(\omega) \geq 0$ for every $\omega \in \Omega$ then we can call f 'positive' or more precisely non-negative. If however $f(\omega) \geq 0$ for $\omega \in \Omega \setminus N$, where $N \in \mathcal{F}$ is a set of zero probability, then we say "$f \geq 0$ almost everywhere, or, (in probabolistic language), almost surely". If $E \in \mathcal{F}$ then we say "$f \geq 0$ on E almost-surely" if there is $N \in \mathcal{F}$, with $N \subseteq E$ and $\mathbb{P}(N) = 0$ and $f(\omega) \geq 0$ $\forall \omega \in E \setminus N$. This kind of thing extends to all kinds of properties of random variables and sets in $\mathcal{F}$. A property, P, is said to hold on $E \in \mathcal{F}$ almost-surely iff there is $N \subseteq E$, $\mathbb{P}(N) = 0$ (so $N \in \mathcal{F}$) and P holds for every $\omega \in E \setminus N$.

Let f, g be random variables. We say $f = g$ almost surely on Ω iff $\exists N \in \mathcal{F}$ with $\mathbb{P}(N) = 0$ and $f(\omega) = g(\omega)$ $\forall \omega \in \Omega \setminus N$. In this case we write $f \sim g$. Of course this idea depends upon the probability measure $\mathbb{P}$. One can show that "$\sim$" is an equivalence relation on the set of random variables. If $f \sim g$ then for any $E \in \mathcal{F}$;

$$\int_E f \, d\mathbb{P} = \int_E g \, d\mathbb{P}$$

because $(f - g) I_E$ is zero, $\mathbb{P}$ a.s. and thus has zero integral

All of this emphasises the fact that sets of zero (probability) measure are negligible in integration theory. Now you might think that a subset of a 'negligible set' should also be negligible. However, there are examples of sets which are subsets of a set of zero measure but are not themselves measurable! This is because the σ-field that the (probability) measure is defined on is not as large as it could be(+). So, we could extend our $\mathbb{P}$ so that $\mathbb{P}(A) = 0$ for every $A \subseteq N$, where $N \in \mathcal{F}$ and $\mathbb{P}(N) = 0$. But is $\mathbb{P}$ still a measure on a σ-field?

Theorem

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $\mathcal{F}^*$ be the collection of all $E \subseteq \Omega$ for which there are sets $H, J \in \mathcal{F}$ with $H \subseteq E \subseteq J$ and $\mathbb{P}(J \setminus H) = 0$. Define $\mathbb{P}(E) = \mathbb{P}(H)$ in this case. Then $\mathcal{F}^*$ is a σ-algebra and $\mathbb{P}$ is a probability measure on $\mathcal{F}^*$. So $(\Omega, \mathcal{F}^*, \mathbb{P})$ is a complete probability space.

This extended version of $\mathbb{P}$ is called 'complete', because all subsets of $\mathbb{P}$-null sets are themselves $\mathbb{P}$-null, and $\mathcal{F}^*$ is called the $\mathbb{P}$-completion of $\mathcal{F}$. The theorem tells us that any $\mathbb{P}$ can be "completed".

(+) Well, this is one possibility.

Pf

(i) If $E \in \mathcal{F}$ then take $H = E = J$ and see that $E \in \mathcal{F}^*$ and that $\mathbb{P}(E)$ remains unchanged.

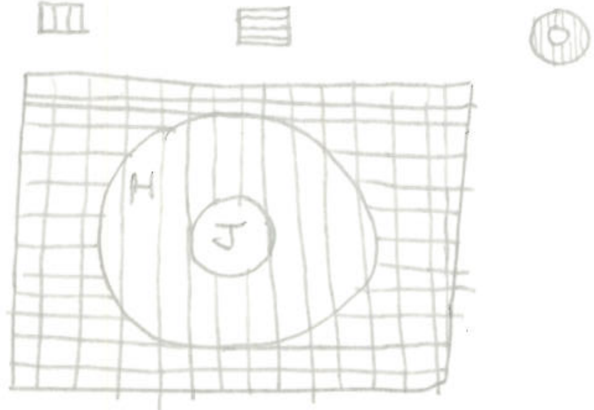
(ii) So $\Omega \in \mathcal{F}^*$

(iii) If $E \in \mathcal{F}^*$ there and H, J in $\mathcal{F}$ with $J \subseteq E \subseteq H$ so $\Omega\setminus H \subseteq \Omega\setminus E \subseteq \Omega\setminus J$ and $\Omega\setminus H, \Omega\setminus J$ lie in $\mathcal{F}$.

But, $\mathbb{P}((\Omega\setminus J)\setminus(\Omega\setminus H)) = \mathbb{P}(H\setminus J) = 0$,

because H, J can be chosen that way and

$$(\Omega\setminus J)\setminus(\Omega\setminus H) = H\setminus J$$



(iii) If $E_i \in \mathcal{F}^*$ $i \in \mathbb{N}$ and $\exists\, J_i \subseteq E_i \subseteq H_i$ with $\mathbb{P}(H_i\setminus J_i) = 0$. Let $J = \bigcup_i J_i$, $E = \bigcup_i E_i$, $H = \bigcup_i H_i$, then $J \in \mathcal{F}$ and $H \in \mathcal{F}$, also $\mathbb{P}(H\setminus J) \le \mathbb{P}\left(\bigcup_i (H_i\setminus J_i)\right)$

because,

$$H\setminus J = \bigcup_i H_i \setminus \bigcup_j J_j \subseteq \bigcup_i (H_i\setminus J_i),$$

$$\boxed{\begin{aligned}\bigcup_i H_i \setminus \bigcup_j J_j &= \bigcup_i \Big(H_i \setminus \bigcup_j J_j\Big)\\ &\subseteq \bigcup_i (H_i\setminus J_i).\end{aligned}}$$

and

$$\mathbb{P}\Big(\bigcup_i H_i\setminus J_i\Big) \le \sum_i \mathbb{P}(H_i\setminus J_i) = 0, \text{ if } \mathbb{P}(H_i\setminus J_i) = 0$$

$\forall i$. So $\bigcup_i E_i \in \mathcal{F}^*$. Suppose now that $E \in \mathcal{F}^*$ and $J \subseteq E \subseteq H$, $J, H \in \mathcal{F}$ with $\mathbb{P}(H\setminus J) = 0$, and, also

$J_1 \subseteq E \subseteq H_1$, $J_1, H_1 \in \mathcal{F}$ with $\mathbb{P}(H_1 \setminus J_1) = 0$, on the face of it we have two possible values for $\mathbb{P}(E)$, $\mathbb{P}(J)$ and $\mathbb{P}(J_1)$(†). We show they are equal: as

$$J \setminus J_1 \subseteq H_1 \setminus J_1 \quad \text{and} \quad \mathbb{P}(H_1 \setminus J_1) = 0$$

then $\mathbb{P}(J \setminus J_1) = 0$. Similarly,

$$J_1 \setminus J \subseteq H \setminus J \quad \text{and} \quad \mathbb{P}(H \setminus J) = 0$$

so $\mathbb{P}(J_1 \setminus J) = 0$. Now,

$$J = J \setminus J_1 \cup J \cap J_1$$

and $J_1 = J_1 \setminus J \cup J \cap J_1$,

so $\mathbb{P}(J) = \mathbb{P}(J \setminus J_1) + \mathbb{P}(J \cap J_1)$

and $\mathbb{P}(J_1) = \mathbb{P}(J_1 \setminus J) + \mathbb{P}(J \cap J_1)$

and $\mathbb{P}(J) = \mathbb{P}(J_1)$. So $\mathbb{P}$ is well defined on $\mathcal{F}^*$.

That $\mathbb{P}$ is a countably additive set function on $\mathcal{F}^*$ is "easy": let $(E_i) \subset \mathcal{F}^*$ with $J_i \subseteq E_i \subseteq H_i$, $\mathbb{P}(H_i \setminus J_i) = 0$, and $E_i \cap E_j = \emptyset$ if $i \neq j$. Well, (J_i) is a disjoint family in $\mathcal{F}$ with $\mathbb{P}(\bigcup_i J_i) = \sum_i \mathbb{P}(J_i)$ (because $\mathbb{P}$ is a measure on $\mathcal{F}$) and $\bigcup_i J_i \subseteq \bigcup_i E_i \subseteq \bigcup_i H_i$ while

$$\mathbb{P}\left(\bigcup_i H_i \setminus \bigcup_i J_i\right) \leq \mathbb{P}\left(\bigcup_i (H_i \setminus J_i)\right) \leq \sum_i \mathbb{P}(H_i \setminus J_i) = 0.$$

So, from the "definitions",

$$\mathbb{P}\left(\bigcup_i E_i\right) = \mathbb{P}\left(\bigcup_i J_i\right) = \sum_i \mathbb{P}(J_i) = \sum_i \mathbb{P}(E_i).$$

□

(†) There are the same as $\mathbb{P}(H)$ and $\mathbb{P}(H_1)$.

Extending the idea of measurable function:

Measurable functions which agree on a set of probability 1 are 'identical' so far as integration theory is concerned. This reinforces the idea that one need only define a function on a set of probability one "to do integration theory". In fact this can be made precise. Define a function, f, to be "measurable on Ω" if there is a set $E \in \mathcal{F}$ with $\mathbb{P}(E) = 1$ and f is defined on E while $f^{-1}(B) \cap E \in \mathcal{F}$ for every Borel Set, $B \subseteq \mathbb{R}$. First of all, it is obvious that any function measurable in the usual sense is 'measurable on Ω' indeed any choice of E leaves the function satisfying 'the conditions'. Moreover, if f is measurable on Ω and we define f to be 0 on $\Omega \setminus E$ then we have a measurable function in the original sense(+). Should $\mathbb{P}$ be complete then we can define f on $\Omega \setminus E$ in any way that we wish (note that each point in $\Omega \setminus E$ – i.e. the singleton containing the point – lies in $\mathcal{F}$ if $\mathbb{P}$ is complete). But the integral of f over any $F \in \mathcal{F}$ is independent of how f might be defined on $\Omega \setminus E$. So we don't need to specify how f is to be defined on $\Omega \setminus E$, so long as integration is our only motive.

(+) Consider $f^{-1}(B)$, $B \subseteq \mathbb{R}$ a Borel set, for f defined in this way.

Then $f^{-1}(B) = f^{-1}(B) \cap E \;\cup\; f^{-1}(B) \cap (\Omega \setminus E)$, the first set on the right side is in $\mathcal{F}$ and the second is $\emptyset$ if $0 \notin B$ and $\Omega \setminus E$ if $0 \in B$ in either event that union lies in $\mathcal{F}$.

It is all very well to define a function, $f : \Omega \to [-\infty, \infty]$ as 'measurable on Ω' but we are used to adding, multiplying, and so on, so is there a sensible way of adding such functions? Obviously, we could extend any pair to all of Ω in the manner described above (making them zero off of their respective E's) and then just add them as usual.

One has to object that this runs against the spirit of the definition (if you're going to do this why bother ----?). But observe if f_1, f_2 are measurable on Ω, with sets E_1, E_2 of probability 1 for which $f_i^{-1}(B) \cap E_i \in \mathcal{F}$, $i = 1, 2$, then

$$E_1 = E_1 \cap E_2 \cup E_1 \setminus E_2,$$

and

$$E_1 \setminus E_2 \subseteq \Omega \setminus E_2 . \qquad \text{So } \mathbb{P}(E_1 \setminus E_2) \leq \mathbb{P}(\Omega \setminus E_2) = 0$$

and therefore $1 = \mathbb{P}(E_1) = \mathbb{P}(E_1 \cap E_2)$, obviously $\mathbb{P}(\Omega \setminus (E_1 \cap E_2)) = 0$. If we define $f_1 + f_2$ as their pointwise sum on $E_1 \cap E_2$ we need only show that $(f_1+f_2)^{-1}(B) \cap (E_1 \cap E_2) \in \mathcal{F}$ for each Borel set $B \subseteq \mathbb{R}$ and we would have a reasonable candidate for $f_1 + f_2$, it would be 'measurable on Ω': Extend f_1 and f_2 to all of Ω by setting them to be zero on $\Omega \setminus E_1$ and $\Omega \setminus E_2$ respectively. We <u>know</u> that "$f_1 + f_2$" is measurable, that $f_1 + f_2$ "extended" is identical with $f_1 + f_2$ on $E_1 \cap E_2$, certainly then $(f_1+f_2)^{-1}(B) \cap (E_1 \cap E_2) \in \mathcal{F}$. So our definition of sum 'works'. Notice also, that if f is measurable on Ω with $f^{-1}(B) \cap E \in \mathcal{F}$, $\forall B$, and F is another set of probability 1, then $f^{-1}(B) \cap E \cap F \in \mathcal{F}$ and $\mathbb{P}(E \cap F) = 1$. There is an overwhelming sense that f defined on E and f defined on $E \cap F$ are but <u>versions</u> of f. This is made precise by using the equivalence classes determined by the relation $f \sim g \iff f = g$ $\mathbb{P}$ a.s. It is these <u>classes</u> that become the objects of our study. We think of them as functions defined only on a set of probability 1, moreover for <u>any</u> set of probability 1 there is a 'version of f' defined on it.