Homework Solutions

Lemma 3.4. Let $\mathcal{F}$ denote the intersection $\cap_\alpha \mathcal{F}_\alpha$. Since each $\mathcal{F}_\alpha$ is a $\sigma$-algebra, we have that $\emptyset \in \mathcal{F}_\alpha$ for all $\alpha$. Thus, $\emptyset \in \mathcal{F}$ as well (because it is the intersection). Similarly, if $A \in \mathcal{F}$, then $A \in \mathcal{F}_\alpha$ for all $\alpha$ and $A^c \in \mathcal{F}_\alpha$ for all $\alpha$ and hence $A^c \in \mathcal{F}$. The same reasoning works for $\cup_\alpha A_n$ where $A_n \in \mathcal{F}$ for all $n \geq 1$. Indeed, $A_n \in \mathcal{F}_\alpha$ for all $\alpha$ (and all $n \geq 1$) and hence $\cup_\alpha A_n \in \mathcal{F}_\alpha$ for all $\alpha$ and consequently $\cup_\alpha A_n \in \mathcal{F}$.

Lemma 3.11(i). If $A_n$ are increasing then let $B_n = A_n \setminus A_{n-1}$. These are disjoint sets and $A_N = \cup_{n=1}^N B_n$. But then

$$\mu(A_N) = \mu(\cup_{n=1}^N B_n) = \sum_{n=1}^N \mu(B_n).$$

Also, note that $\cup_\alpha A_n = \cup_\alpha B_n$. Taking $n \to \infty$ and using $\sigma$-additivity shows

$$\lim_{N \to \infty} \mu(A_N) = \sum_{n=1}^\infty \mu(B_n) = \mu(\cup_\alpha B_n) = \mu(\cup_\alpha A_n).$$

Conversely, if say we did not know we have $\sigma$-additivity but knew we had additivity and continuity under increasing unions. Then consider $B_n$ disjoint. Sets $A_N = \cup_{n=1}^N B_n$ are increasing to $\cup_\alpha B_n$. We now have

$$\mu(\cup_\alpha B_n) = \mu(\cup_\alpha A_n) = \lim_{N \to \infty} \mu(A_N) = \lim_{N \to \infty} \sum_{n=1}^N \mu(B_n) = \sum_{n=1}^\infty \mu(B_n).$$

The second-last equality comes from the assumed additivity.

Lemma 3.11(ii). This follows from the above by using complements. However, one issue is that if $\mu(A_n) = \infty$ then we cannot deduce what $\mu(A_n^c)$ is. Hence the need for finiteness for this part of the argument. Namely, that if $\mu$ is finite and we have additivity and continuity under decreasing intersections, then we have $\sigma$-additivity.

Lemma 3.9. $\mu(A \cup B) = \mu(A) + \mu(B \setminus A)$ because of additivity and the two sets on the right-hand side being disjoint and their union making up the set on the left-hand side. Similarly, $\mu(B) = \mu(B \setminus A) + \mu(A \cap B) \geq \mu(B \setminus A)$. Together these give

$$\mu(A \cup B) \leq \mu(A) + \mu(B).$$
(In fact, from the above we get that if \( \mu(A \cap B) < \infty \) then \( \mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B) \). Of course, if it is infinite then \( \mu(A \cup B) = \infty \) too.)

Induction gives

\[
\mu(\bigcup_{n=1}^{N} A_n) \leq \sum_{n=1}^{N} \mu(A_n).
\]

Since \( \bigcup_{n=1}^{N} A_n \) increases to \( \bigcup A_n \), a \( \sigma \)-additive \( \mu \) would give

\[
\mu(\bigcup_{n=1}^{\infty} A_n) = \lim_{n \to \infty} \mu(\bigcup_{n=1}^{N} A_n) \leq \lim_{n \to \infty} \sum_{n=1}^{N} \mu(A_n) = \sum_{n} \mu(A_n).
\]

**Borel \( \sigma \)-algebra is generated by intervals:** Let \( B \) be the Borel \( \sigma \)-algebra. Let \( F \) be the \( \sigma \)-algebra generated by finite unions of intervals of the form \((a, b]\). Since finite unions of intervals are clearly \( B \)-measurable we know that \( B \) must contain \( F \). On the other hand, any interval \((a, b]\) is an intersection of intervals \((a, b-1/n]\), which are all open. Hence, intervals of the form \((a, b]\) are all in \( F \). But any open set can be written as a (countable) union of such intervals. Hence, open sets are all in \( F \) and so is the \( \sigma \)-algebra generated by them. But this is \( B \). In other words we also have \( B \subset F \).

3.2. Let \( \Omega = (0, 1] \). For \( n \geq 1 \) let \( F_n \) be made up of \( \emptyset \) and any unions of intervals of the form \((i/2^n, (i+1)/2^n]\), with \( i \in \{0, 1, \ldots, 2^n - 1\} \). This is an algebra (CHECK!). But since it is made up of finitely many sets, it is also a \( \sigma \)-algebra. The union of \( F_n \) clearly does not have all open sets, since not all open sets are in SOME \( F_n \). On the other hand, any interval of the form \((a, b]\) is a union of intervals \(((i_n+1)/2^n, j_n/2^n]\), where \( i_n \) is the largest integer such that \( i/2^n < a \leq (i+1)/2^n \) and \( j_n \) is the smallest integer such that \( j_n/2^n < b \leq (j_n+1)/2^n \). Note that \((i_n+1)/2^n \leq 1/2n \) and \( b - j_n/2^n \leq 1/2n \). Furthermore, \( i_n \) is nondecreasing and \( j_n \) nonincreasing (CHECK!). Therefore, interval \(((i_n+1)/2^n, j_n/2^n]\) (which is in \( F_n \)) increases to \((a, b]\). This proves that intervals \((a, b]\) are all in the \( \sigma \)-algebra generated by \( \bigcup_n F_n \). But then so are open sets. Hence, the \( \sigma \)-algebra generated by \( \bigcup_n F_n \) is not \( \bigcup_n F_n \) itself. It is bigger. In fact, the proof shows that it is exactly the Borel \( \sigma \)-algebra on \((0, 1]\).

An easier example: Let \( \Omega = \mathbb{N} \) and let \( F_n \) be the collection of subsets of \( \{1, \ldots, n\} \) and their complements in \( \mathbb{N} \) (i.e. subsets of the form \( A \cup \{n+\})
1, n + 2, \ldots \} \) where \( A \) is a subset of \( \{1, \ldots, n\} \). Since \( \mathcal{F}_n \) is an algebra (CHECK!) with finitely many sets in it, it is also a \( \sigma \)-algebra. This sequence is increasing, and the \( \sigma \)-algebra generated by the union must have in it the set of even numbers (because that is the countable union of sets \( \{2n\} \in \mathcal{F}_{2n} \)). But the set of all even numbers is not in any of \( \mathcal{F}_n \). Hence the union is not the same as the \( \sigma \)-algebra generated by the union.

3.9. If \( F(a) = \mu((-\infty, a]) \) then \( F(a) \) is nondecreasing because if \( a < b \) then \((-\infty, a] \subset (-\infty, b] \) and thus the \( \mu \)-measures are ordered the same way and as a result \( F(a) \leq F(b) \). It also increases to 1 as \( a \to \infty \) and decreases to 0 as \( a \to -\infty \) and is right-continuous, all for the same reason: \( \sigma \)-additivity of \( \mu \) implies its continuity under monotone limit (decreasing limits are OK because \( \mu \) is a probability measure and hence is finite). Right-continuity comes from the fact that if \( a_n \) decreases to \( a \) then \((-\infty, a_n] \) decrease to \((-\infty, a] \) and hence the measures converge, in other words \( F(a_n) \) converges to \( F(a) \). Convergence to 1 and to 0 comes from the fact that \((-\infty, a] \) decrease to \( \emptyset \) as \( a \to -\infty \) and increase to all of \( \mathbb{R} \) as \( a \to \infty \).

Conversely, if \( F(a) \) is a CDF, then we can define \( \mu((a, b]) = F(b) - F(a) \). We can then define the measure of a finite intersection of such intervals. This defines an additive (probability) measure on an algebra. We can extend this measure uniquely to the \( \sigma \)-algebra generated by this algebra (which is nothing other than the Borel \( \sigma \)-algebra) by means of Carathéodory’s theorem. But for this we need to show that \( \mu \) is \( \sigma \)-additive on the algebra. This can be shown in exactly the same way as was shown for the Lebesgue measure, except that this time we use \( F(b) - F(a) \) instead of \( b - a \) to measure an interval. One thing that was needed in that construction is that if \( a \) is given, then we can find a \( b > a \) close to \( a \) to make \( F(b) - F(a) \) as small as we wish. This is where right-continuity comes in. The responsible student should fill in the details (e.g. adapting the proof of the Lebesgue measure case).

3.10. First, let us show that it makes sense to talk about \( a + A \) and \( cA \). Namely, if \( a, c \in \mathbb{R} \) are numbers, then the functions \( x \mapsto x - a \) and \( x \mapsto x/c \) are continuous. We have seen in class (albeit in lectures pertaining to chapter 4!) that these functions are then measurable. In particular, if \( A \) is a Borel set, then \( a + A \) (which is the inverse image of \( A \) under \( x \mapsto x - a \)) and \( cA \) (inverse image of \( x \mapsto x/c \)) are also Borel sets. To do this more directly one can consider \( \mathcal{C} = \{ A \in \mathcal{B} : a + A \in \mathcal{B} \} \) and observe that this clearly contains open sets and is a monotone class. Hence it contains \( \mathcal{B} \) (the
\(\sigma\)-algebra generated by the open sets), but then this means it equals \(B\). In other words, for any \(A \in B\) we have \(a + A \in B\). The same can be done for \(cA\).

Now that we have established that shifting and scaling leave the sets measurable we want to prove that the Lebesgue measure is invariant under shifts. For this, fix an \(a \in \mathbb{R}\) and define the measure \(\mu\) that assigns values \(\mu(A) = m(a + A)\). Here, \(m\) is Lebesgue measure. Note that if \(A\) is an interval then \(\mu(A)\) is in fact the same thing as \(m(A)\) (both are equal to the length of the interval). So \(\mu\) and \(m\) coincide on intervals and hence also on finite disjoint unions of intervals. Because both measures are clearly \(\sigma\)-finite, we deduce that they must match on the whole Borel \(\sigma\)-algebra (by the uniqueness part in Carathéodory’s theorem). This says that \(m(a + A) = m(A)\), which is what we wanted to prove.

For the scale invariance we do the same as above: define a measure \(\nu(A) = \alpha m_1(\alpha^{-1}A)\) on the Borel \(\sigma\)-algebra of \([0, \alpha]\). We want to show that \(\nu = m_\alpha\). But for this, it is enough to check that the two measures match on intervals. This in turn is pretty clear: \(\nu((a, b]) = \alpha m_1((a/\alpha, b/\alpha]) = b - a = m_\alpha((a, b])\).

By the uniqueness of the extension we deduce that the two measures actually match on the whole \(B\).

3.11. First, let us assume that \(\mu((\alpha, \beta]) < \infty\) for some \(\alpha < \beta\). By shift-invariance we have \(\mu((\alpha, \beta]) = \mu((0, \beta - \alpha])\). So we can assume \(\alpha = 0\) (and rename \(\beta - \alpha > 0\) as \(\beta > 0\)). Abbreviate \(f(t) = \mu((0, t])\). Notice that for \(t, s > 0\)

\[
f(t+s) = \mu((0, t+s]) = \mu((0, t])+\mu((t, t+s]) = \mu((0, t])+\mu((0, s]) = f(t)+f(s).
\]

The first equation came by additivity and the second by the assumed shift-invariance of \(\mu\). By induction we have that for any \(n \geq 1\) and \(s > 0\) we have \(f(ns) = nf(s)\). One consequence of this is that \(\mu((0, n\beta]) = n\mu((0, \beta]) < \infty\) for all \(n\). Hence, for any \(a < b\) we have \(\mu((a, b]) = \mu((0, b-a]) \leq \mu((0, n\beta]) < \infty\), where we choose \(n\) large enough to have \(b - a \leq n\beta\). As a result, we see that we could in fact choose \(\beta = 1\).

Substituting \(s\) in the above display with \(s/n\) we get \(f(s) = nf(s/n)\) or equivalently \(f(s/n) = (1/n)f(s)\). These last two facts say that if \(t = m/n\) with \(m, n \in \mathbb{N}\), and if \(s = 1\), then \(f(t) = f(m/n) = mf(1/n) = mf(1)/n = tf(1)\). Furthermore, if \(t > 0\) is irrational, then we can find a rational sequence
Then $(0, t_n]$ decreases to $(0, t]$ and since we have shown that $\mu((a, b]) < \infty$ for all $a < b$, we have that $\mu((0, t + 1]) < \infty$ and $\sigma$-additivity of $\mu$ implies that $\mu((0, t_n])$ decreases to $\mu((0, t])$. Thus

$$f(t) = \lim_{n \to \infty} f(t_n) = \lim_{n \to \infty} t_n f(1) = t f(1).$$

We have thus shown that

$$\mu((0, t]) = \mu((0, 1]) t, \quad \text{for all } t > 0.$$  

Using shift-invariance of $\mu$ and that $m((a, b]) = b - a$ we have

$$\mu((a, b]) = \mu((0, b - a]) = \mu((0, 1])(b - a) = \mu((0, 1]) m((a, b]), \quad \text{for all } b > a.$$  

This says that the measures $\mu$ and $\mu((0, 1]) m$ match on semi-open intervals. Since they are both $\sigma$-additive and $\sigma$-finite, uniqueness in Carathéodory’s theorem says they must match on the whole Borel $\sigma$-algebra. This answers the question with $c = 1/\mu((0, 1])$. Note that all this makes sense even if $\mu((0, 1]) = 0$. In this case, then $c = \infty$ and $\mu$ assigns 0 to everything.

It remains to dismiss the case when $\mu((a, b]) = \infty$ for all $a < b$. It may seem at first that $\sigma$-finiteness alone should suffice. However, consider the measure $\nu$ that assigns to a Borel set $A$ the number of rationals in that set. You can quickly check that this is a $\sigma$-additive nonnegative measure. It is also $\sigma$-finite because we can number the rationals as say $\{q_n : n \geq 1\}$ and then consider the events $A_n = \left((-n, n) \setminus \mathbb{Q}\right) \cup \{q_1, \ldots, q_n\}$. It is clear that $A_n$ increases to $\mathbb{R}$ and that $\nu(A_n) = n < \infty$. But even though measure $\nu$ is $\sigma$-additive, it assigns an infinite value to all intervals. It is not shift-invariant, though. So we must use shift-invariance crucially in our argument.

The actual argument requires knowing a bit more probability, namely knowing about product measures and the Radon-Nikodým derivative. We will thus revisit this exercise when we learn more about these.

### 3.13 Define $\text{supp}(\mu)$, the support of $\mu$, to be the intersection of all closed sets $C$ such that $\mu(C^c) = 0$. Since the whole space $\Omega$ is one such set, the intersection is not empty. Since any intersection of closed sets is always a closed set, the support supp($\mu$) is a closed set and is hence measurable. Next, note that $x \not\in \text{supp}(\mu)$ is equivalent to saying that $x$ is in the union of all $C^c$.
such that \( \mu(C^c) = 0 \). In other words, this is equivalent to the existence of an open set \( U \) such that \( \mu(U) = 0 \).

**Comment:** The problem in the textbook asks for a bit more than the above. It says that the support of \( \mu \) is THE smallest closed set whose complement has measure 0. That is, it is asking us to also show that the measure of the complement of the support is 0. This is not a trivial matter. Let us denote the complement of the support by \( A \). Then we know that \( A \) is the union of all open sets \( U \) with \( \mu(U) = 0 \). But since this union can be uncountable, in a general topological space, we cannot deduce that \( \mu(A) = 0 \). (And there are in fact counter-examples.) However, if the space is also second-countable (e.g. metric and separable), then there is a countable collection \( O_n \) of open sets such that each open set \( U \) can be written as a countable union of a subcollection of such sets. If \( U \) happens to also satisfy \( \mu(U) = 0 \), then every \( O_n \) in the subcollection will also have \( \mu(O_n) \) (because \( O_n \) is inside of \( U \)). This shows that \( A \) can in fact be written as a countable union of of open sets \( O_n \) with \( \mu(O_n) = 0 \) for each \( n \). Now this then guarantees that \( \mu(A) = 0 \).

3.14. Let \( A_m \in \mathcal{M} \) be increasing. Since \( A_m \in \mathcal{M} \) we have for each \( n \) a set \( B_{m,n} \in \mathcal{A} \) such that \( \mu(A_m \Delta B_{m,n}) \leq 1/(2^m n) \). For a given \( n \) let \( B = \bigcup_m B_{m,n} \). Write

\[
\mu(A \Delta B) = \mu(\bigcup_m A_m \Delta \bigcup_m B_{m,n}) \leq \mu \left( \bigcup_m (A_m \Delta B_{m,n}) \right)
\]

\[
\leq \sum_{m \geq 1} \mu(A_m \Delta B_{m,n}) \leq \frac{1}{n} \sum_{m=1}^{\infty} \frac{1}{2^m} = \frac{1}{n}.
\]

It seems like we have thus shown that \( A \in \mathcal{M} \). In other words, \( \mathcal{M} \) is closed under increasing unions. One caveat, though: we have not shown that \( B \in \mathcal{A} \). This would be the case if \( \mathcal{A} \) were a \( \sigma \)-algebra. But \( \mathcal{A} \) is a monotone class, and so it is only closed under increasing unions, not any arbitrary unions. So it is not clear that the way we chose \( B \) keeps it in \( \mathcal{A} \). Also, we have not used the fact that \( A_m \)'s are increasing.

So we will proceed a little differently from above. First, let \( A = \cup_{m \geq 1} A_m \). We want to show that \( A \in \mathcal{M} \). So fix any \( n \). We want to find a \( B \in \mathcal{A} \) such that \( \mu(A \Delta B) \leq 1/n \). We have that \( \mu(A_m) \) increases to \( \mu(A) \). So there exists an \( m \) such that \( \mu(A \setminus A_m) = \mu(A) - \mu(A_m) \leq 1/(2n) \). And since \( A_m \) itself is in \( \mathcal{M} \), there exists a \( B \in \mathcal{A} \) such that \( \mu(A_m \Delta B) \leq 1/(2n) \). But now \( A \Delta B \subset (A_m \Delta B) \cup (A \setminus A_m) \) (CHECK!). Hence

\[
\mu(A \Delta B) \leq \mu(A_m \Delta B) + \mu(A \setminus A_m) \leq 1/(2n) + 1/(2n) = 1/n,
\]
as desired. Closure under decreasing intersections works similarly (but the student should CHECK).

Note that this time around, we did not use the fact that $\mathcal{A}$ is a monotone class! $\mathcal{A}$ could be any collection of measurable sets, and $\mathcal{M}$ would still be a monotone class. The point of the exercise is that if you have a collection of sets $\mathcal{A}_0$ with which you can approximate the measure of sets in an algebra $\mathcal{A}$, in the sense that for each $A \in \mathcal{A}$ and each $\varepsilon > 0$ there exists a $B \in \mathcal{A}_0$ such that $\mu(A \Delta B) < \varepsilon$, then you can also approximate the sets $A \in \sigma(\mathcal{A})$ by sets $B \in \mathcal{A}_0$, in that same sense.

3.17. (1) is straightforward. (2) is a direct use of the axiom of choice: you have a family of nonempty sets and you want to choose one representative from each set. For (3) say that the intersection is not empty. Then there exist $s, t \in (0, 1]$ such that $e^{2\pi i s}, e^{2\pi i t} \in \Lambda$, and $e^{2\pi i a} e^{2\pi i s} = e^{2\pi i b} e^{2\pi i t}$. But then $e^{2\pi i (t - s)} = e^{2\pi i (a - b)}$. Since $a - b$ is rational there exists an integer $n$ such that $n(a - b)$ is a whole number. Then $1 = e^{2\pi i n (a - b)} = e^{2\pi i n (t - s)}$. This implies that $n(t - s)$ is a whole number and $t - s$ is rational. But then $e^{2\pi i t} \sim e^{2\pi i s}$ and they belong to the same equivalence class. This forces $t = s$ for otherwise $e^{2\pi i s}$ and $e^{2\pi i t}$ cannot represent two distinct elements in $\Lambda$. But now if $t = s$, then we also have $\alpha = \beta$, which was assumed not to be the case.

(5) really follows from (4), for if $\Lambda$ were measurable, then we would be able to compute $\mu(\Lambda)$. To show (4) let us first prove the statement in the hint: take element $e^{it}$ with $t \in (0, 2\pi]$. Let $s \in \Lambda$ be the representative of the equivalence class of $t$. Set $\alpha = t - s$. Since $t \sim s$, we have $\alpha \in \mathbb{Q}$. But then $e^{it} = e^{ia} e^{is}$, i.e. $e^{it} \in \Lambda_\alpha$, and the claim in the hint is proved. (One little detail: it could be that $t < s$ and thus $\alpha$ is negative. In this case, use $\alpha' = 2\pi + \alpha$, which is then in $(0, 2\pi]$ and we have $e^{it} = e^{i\alpha'} e^{is} \in \Lambda_{\alpha'}$.)

Next, we have just shown that $\Lambda_\alpha$ are all disjoint. Furthermore, we have already shown that Lebesgue measure is shift-invariant and since $\Lambda_\alpha$ is essentially a shift of $\Lambda$ we have $\mu(\Lambda_\alpha) = \mu(\Lambda)$ for all $\alpha$. (This actually needs proof, but the proof goes along the same lines as that of problem 3.10.) But then $\sigma$-additivity says that

$$2\pi = \mu(S^1) = \sum_{\alpha \in (0, 2\pi] \cap \mathbb{Q}} \mu(\Lambda_\alpha) = \sum_{\alpha \in (0, 2\pi] \cap \mathbb{Q}} \mu(\Lambda).$$
Either $\mu(\Lambda) = 0$ and then we would have $2\pi = 0$, or $\mu(\Lambda) > 0$ and then $2\pi = \infty$. Therefore, it must be the case that $\mu(\Lambda)$ is not defined, which means $\Lambda$ is not measurable.
4.1 That the collection is a $\sigma$-algebra follows by checking all the necessary properties in the definition of a $\sigma$-algebra (and using the fact that $X^{-1}(A \cup B) = X^{-1}(A) \cup X^{-1}(B)$ and $X^{-1}(A^c) = (X^{-1}(A))^c$). That it is the smallest $\sigma$-algebra on which $X$ is measurable is also clear: $X$ is clearly measurable with respect to $\sigma(X)$ (because for every $A \in \mathcal{A}$ we have $B \in \sigma(X)$ by the definition of $\sigma(X)$!) and if $X$ is measurable with respect to some $\sigma$-algebra $\mathcal{F}$, then for every $A \in \mathcal{A}$ we must have $X^{-1}(A) \in \mathcal{F}$, but this says that $\sigma(X) \subset \mathcal{F}$!

4.6. First observe that we can approximate $1_{(a,b]}$ by a sequence of continuous functions, e.g. $f_n(x) = 0$ for $x \leq a$ or $x \geq b + 1/n$, $f_n(x) = 1$ for $a + 1/n \leq x \leq b$, and interpolating linearly on $(a, a + 1/n)$ and $(b, b + 1/n)$. Next, we can approximate the indicator function of $A = \prod_{i=1}^d (a_i, b_i]$ by functions $f_n(x_1, \ldots, x_d) = \prod_{i=1}^d f_{n,i}(x_i)$, where $f_{n,i}$ approximates $(a_i, b_i]$. We can thus approximate indicators of finite disjoint unions of such boxes, by continuous functions. We have that

$$
\mu(|f_n - 1_A| > \varepsilon) \leq \mu(f_n \neq 1_A) \leq \sum_{i=1}^d \mu(f_{n,i} \neq 1_{(a_i, b_i]})
$$

$$
\leq \sum_{i=1}^d \left( \mu([a, a + 1/n]) + \mu([b, b + 1/n]) \right) \rightarrow 0.
$$

(The last limit is due to $\sigma$-additivity.) We have that for each $n$, $\int f_n \, d\mu = \int f_n \, d\nu$. Take $n \to \infty$. The approximating functions are all bounded and so bounded convergence allows us to slip the limit under the integrals and we get that $\mu$ and $\nu$ match on finite unions of semi-open boxes. These form an algebra and since both measures are assumed to be $\sigma$-finite, the uniqueness in Carathéodory’s theorem implies the two measures must be equal.

Counter-example to Fatou. Consider Lebesgue measure on $[0, 1]$. Let $f_n = -n 1_{(0, 1/n]}$. Then $\lim f_n(x) = 0$ for all $x$. But $\int f_n(x) \, dx = -1$ for all $n$. Fatou’s lemma fails here, since it says that $\int 0 \, dx \leq -1$, which is false.

4.16. Say $X_n \to X$ in probability. If we show that there exists a subsequence that converges almost surely, then we can repeat this argument starting with any subsequence, to show that that subsequence has a further subsequence converging almost surely.
For a given \( m \) we have that \( P(|X_n - X| > 1/m) \to 0 \) as \( n \to \infty \). Thus, there exists an \( n_m \) such that \( P(|X_{nm} - X| > 1/m) < 1/2^m \). Now we claim that with probability one there exists an \( m_0 \) such that \( |X_{nm} - X| \leq 1/m \) for all \( m \geq m_0 \). This then would imply that \( X_{nm} \to X \) almost surely. To prove our claim we instead show that the probability of its complement is zero. For this we write
\[
\Pr(\forall m_0 \geq 1 \exists m \geq m_0 : |X_{nm} - X| > 1/m) = \Pr \left( \bigcap_{m_0 \geq 1} \bigcup_{m \geq m_0} \{|X_{nm} - X| > 1/m\} \right)
\]
\[
= \lim_{m_0 \to \infty} \Pr \left( \bigcup_{m \geq m_0} \{|X_{nm} - X| > 1/m\} \right) \leq \lim_{m_0 \to \infty} \sum_{m \geq m_0} P(|X_{nm} - X| > 1/m)
\]
\[
\leq \lim_{m_0 \to \infty} \sum_{m \geq m_0} \frac{1}{2^m} = 0.
\]

4.18. There are several ways to do this. I will do it in several steps. First, observe that any function \( f \) can be written as \( f = f^+ - f^- \) and so it is enough to approximate nonnegative functions. Second, notice that if \( f \geq 0 \) is in \( L^p \) and we define \( f_n = f \wedge n \) then
\[
\int |f - f_n|^p d\mu = \int (f - n)^p 1\{f \geq n\} d\mu \leq \int f^p 1\{f \geq n\} d\mu.
\]
The last integral converges to 0 by dominated convergence because \( 1\{f \geq n\} \to 0 \) and \( f^p 1\{f \geq n\} \leq f^p \in L^1 \). The upshot is that we can approximate \( L^p \) functions by bounded ones and so our problem reduces to approximating (in \( L^p \) norm) nonnegative bounded functions by continuous ones. But now that the functions are bounded, we can in fact even do better.

To this end, note that we can approximate nonnegative bounded functions pointwise by simple functions. Indeed, if \( f \leq 0 \) is bounded by \( M \) then for a given \( n \) and \( 0 \leq i \leq nM \) \( A_i = f^{-1}([i/n, (i+1)/n)) \). This is a measurable set and if we let
\[
f_n = \sum_{0 \leq i \leq nM} \frac{i}{n} 1_{A_i}
\]
then it is immediate that \( 0 \leq f - f_n \leq 1/n \) everywhere. In other words, \( f_n \to f \) pointwise and \( f_n \) is a simple function.

Now, it suffices to approximate (in \( L^p \) norm) indicator functions \( 1_A \) by continuous functions. (CHECK that this is enough to deduce that any simple
function can be approximated by continuous functions.) If set $A$ is a box of the form $\prod_{j=1}^{d}(a_{j}, b_{j}]$, then we have seen in Problem 4.6 how to approximate $1_{A}$ pointwise with a continuous function and then dominated convergence implies that this approximation also works in $L^{p}$. This approximation also works for a disjoint union of such boxes, which forms an algebra. For a general measurable set $A$ we want to avoid a messy construction of an approximation and so we rely on an abstract theorem! Let

$$C = \{ A \in \mathcal{B}([0,1]^{d}) : \exists f_{n} \text{ continuous, such that } \|f_{n} - 1_{A}\|_{p} \to 0 \}.$$  

We claim that this is a monotone class. If so, then we just argued that it contains an algebra that generates all of the Borel $\sigma$-algebra. It thus must equal the Borel $\sigma$-algebra, which means that we can approximate any indicator function by continuous functions.

To show that $C$ is a monotone class consider an increasing sequence $A_{n}$ of measurable sets from $C$. Then for each $n$ there exists a continuous function $f_{n}$ such that $\|f_{n} - 1_{A_{n}}\|_{p} \leq 1/n$. Now let $A_{n} = \bigcup_{n}A$ and write

$$\|f_{n} - 1_{A}\|_{p} \leq \|f_{n} - 1_{A_{n}}\|_{p} + \|1_{A_{n}} - 1_{A}\|_{p} \leq 1/n + \|1_{A \setminus A_{n}}\|_{p} = 1/n + \mu(A \setminus A_{n})^{1/p} \to 0 \quad \text{as} \quad n \to \infty.$$  

This shows that the sequence of continuous functions $f_{n}$ approximates $1_{A}$ in $L^{p}$ norm and thus $C$ is closed under increasing unions. Closure under decreasing intersections comes similarly (pretty much word for word, but do CHECK!).

To recap: we have shown that indicator functions can be approximated by continuous functions, therefore simple functions have the same property. But bounded functions can be approximated by simple functions and thus by continuous functions. Nonnegative functions can be approximated by bounded ones and hence by continuous ones. And finally any function can be written as a combination of two nonnegative ones. This proves the claim of the problem.