1. For any $\varepsilon > 0$, one has

$$P(|X_n - X| > \varepsilon) = P(|X_n - c| > \varepsilon) = P(X_n - c < -\varepsilon) + P(X_n - c > \varepsilon)$$

$$\leq F_{X_n}(c - \varepsilon) + 1 - F_{X_n}(c + \varepsilon) \to 0, \quad \text{as } n \to \infty,$$

where we used the fact that $c \pm \varepsilon$ are continuity points of the DF $F_X(t) = 1(t \geq c)$ (this is the step function corresponding to the degenerate distribution $P_X = \varepsilon c$, right?).

2. (a) Start with plotting the DF of $X_k$, which have this form:

$$F_{X_k}(t) = \begin{cases} 
0, & t < -1, \\
1 - \frac{1}{k^2}, & -1 \leq t < k^2, \\
1, & t \geq k^2.
\end{cases}$$

One can easily see that $F_{X_k}(t) = 0$, $t < -1$, for all $k \geq 1$, whereas for $t \geq 1$ one has $F_{X_k}(t) \to 1$ as $k \to \infty$. So, at all points $t \in \mathbb{R}$, the DFs converge to the step function $1(t \geq -1)$, which is the DF of $X \equiv -1$.

(b) It follows from the result of Problem 1 that $X_k \xrightarrow{P} X \equiv -1$ as $k \to \infty$.

But there is no convergence neither in $L^1$ nor in $L^2$. It suffices to show the former is true (as convergence in $L^2$ always implies that in $L^1$, by Lyapunov’s inequality, see lecture slides!). Look: $E |X_k - X| = E(X_k + 1) = 0 \times (1 - k^{-2}) + (k^2 + 1) \times k^{-2} = 1 + k^{-2} \to 1 \neq 0$, $k \to \infty$, no convergence in $L^1$.

Re the a.s. convergence: consider events $A_k := \{X_k = k^2\} \equiv \{X_k \neq -1\}$, $k = 1, 2, \ldots$ Clearly, $\sum_{k \geq 1} P(A_k) = \sum_{k \geq 1} k^{-2} < \infty$, and so by Borel–Cantelli’s lemma (slide 26), $P(A_n, \text{ i.o.}) = 0$. Equivalently, $P(A_n, \text{ f.o.}) = 1$.

For a given $\omega \in \Omega$, denote by $K(\omega)$ the last value of $k$ s.t. $\omega \in A_k$ (putting $K(\omega) := \infty$ if there are infinitely many $k$’s s.t. $\omega \in A_k$). Then the last assertion (that $P(A_n, \text{ f.o.}) = 1$) is equivalent to stating that $P(K(\omega) < \infty) = 1$.

And look what happens: for all $k > K(\omega)$, one has $X_k(\omega) = -1$. That is, on the event $\{K(\omega) < \infty\}$ (which has probability one), one has $X_k \xrightarrow{a.s.} X \equiv -1$.

(c) Look: on the event $\{K(\omega) < \infty\}$, for $n > K(\omega)$, one has

$$Y_n(\omega) = \frac{1}{n} \sum_{k=1}^{n} X_k(\omega) = \frac{1}{n} \sum_{k=1}^{K(\omega)} X_k(\omega) + \frac{1}{n} \sum_{k=K(\omega)+1}^{n} X_k(\omega)$$

$$\leq \frac{1}{n} \sum_{k=1}^{K(\omega)} k^{1/2} K(\omega)(K(\omega)+1)(2K(\omega)+1) \to 0 = \frac{n-K(\omega)}{n} \xrightarrow{n \to \infty} -1$$

as $n \to \infty$. As $P(K(\omega) < \infty) = 1$, that means that $Y_n \xrightarrow{a.s.} Y \equiv -1$. 

1
Re the expectations: as $n \to \infty$,

$$
\mathbb{E} Y_n = \frac{1}{n} \sum_{k=1}^{n} \mathbb{E} X_k = \frac{1}{n} \sum_{k=1}^{n} \left[ -1 \times (1 - k^{-2}) + k^{2} \times k^{-2} \right] = \frac{1}{n} \sum_{k=1}^{n} k^{-2} \to 0.
$$

On the other hand, $\mathbb{E} Y = -1$. So $\mathbb{E} Y_n \not\to \mathbb{E} Y$.

This is not compatible with $Y_n \overset{L^1}{\to} Y$. Indeed, $|\mathbb{E} Y_n - \mathbb{E} Y| = |\mathbb{E} (Y_n - Y)| \leq \mathbb{E} |Y_n - Y|$ (cf. bottom of slide 65), so if the RHS tends to zero (i.e., $Y_n \overset{L^1}{\to} Y$), then one must have $\mathbb{E} Y_n \to \mathbb{E} Y$.

3. (a) Look: as $X_n > 0$, one has

$$
\mathbb{P} \left( |X_n - 0| > \varepsilon \right) = \mathbb{P} (X_n > \varepsilon) = \mathbb{P} (Y_n > n\varepsilon) = \mathbb{P} (Y_1 > n\varepsilon) = 1 - F_{Y_1}(n\varepsilon) \to 0
$$

as $n \to \infty$ since $F_{Y_1}(n\varepsilon) \to 1$ then.

(b) For $\varepsilon > 0$, set $A_n(\varepsilon) := \{|X_n| > \varepsilon\} \equiv \{Y_n > n\varepsilon\}$. Observe that if we show that, $\forall \varepsilon > 0$, one has $\mathbb{P} (A_n(\varepsilon), \text{ i.o.}) = 0$, that would prove that $X_n \overset{a.s.}{\to} 0$.

Indeed, let $B_k := [A_n(1/k), \text{ i.o.}], k = 1, 2, \ldots$ Then $\mathbb{P} \left( \bigcup_{k \geq 1} B_k \right) \leq \sum_{k \geq 1} \mathbb{P} (B_k) = 0$ as well, right? Switching to the complements, we obtain

$$
1 = \mathbb{P} \left( \bigcap_{k \geq 1} B_k^c \right) = \mathbb{P} \left( \bigcap_{k \geq 1} [A_n(1/k), \text{ f.o.}] \right).
$$

Now, on the event $\bigcap_{k \geq 1} [A_n(1/k), \text{ f.o.}]$ (of which the probability is one!), $\forall k \geq 1$ one has $|X_n| > 1/k$ for only finitely many $n$ values. That is, for all large enough $n$ we will have $|X_n| \leq 1/k$. But this is exactly what $X_n \overset{a.s.}{\to} 0$ means!

OK, so we just have to show that, $\forall \varepsilon > 0$, one has $\mathbb{P} (A_n(\varepsilon), \text{ i.o.}) = 0$. By Borel–Cantelli lemma (slide 26), to this end it suffices to show that $\sum_{n \geq 1} \mathbb{P} (A_n(\varepsilon)) < \infty$.

But

$$
\sum_{n \geq 1} \mathbb{P} (A_n(\varepsilon)) = \sum_{n \geq 1} \mathbb{P} (Y_n > n\varepsilon) \leq \frac{1}{\varepsilon} \int_{0}^{\infty} (1 - F_{Y_1}(t)) dt = \frac{1}{\varepsilon} \mathbb{E} Y_1 < \infty, \quad \text{QED.}
$$

Why does the first inequality hold? Partition $(0, \infty)$ into intervals of length $\varepsilon$ and consider the step function $G(t)$ defined by

$$
G(t) := \mathbb{P} (Y_1 > n\varepsilon), \quad t \in ((n-1)\varepsilon, n\varepsilon], \quad n = 1, 2, \ldots
$$

Then $\int_{0}^{\infty} G(t) dt = \sum_{n \geq 1} \varepsilon \mathbb{P} (Y_1 > n\varepsilon)$ and $G(t) \leq 1 - F_{Y_1}(t), t > 0$. Hence that inequality. Draw a picture!

4. Clearly, $X_n \in [0, 1]$ always, so we need to show that, $\forall x \in (0, 1)$, $\mathbb{P} (X_n \leq x) \to$ some DF $F(x)$, as $n \to \infty$. 

2
But look (it would help if you first plot the graph of the function $g(y) := ny - \lfloor ny \rfloor$, $y \in [0, 1]$):

$$
P(X_n \leq x) = P(nY - \lfloor nY \rfloor) = \sum_{k=0}^{n-1} P\left(Y \in \left[\frac{k}{n}, \frac{k+x}{n}\right]\right) = \sum_{k=0}^{n-1} \underbrace{\int_{\frac{k}{n}}^{\frac{k+x}{n}} f(t) \, dt}_{= \frac{x}{n} f(t_k)}
$$

for some $t_k \in \left[\frac{k}{n}, \frac{k+x}{n}\right] \subset \left[\frac{k}{n}, \frac{k+1}{n}\right]$, by the mean value theorem. Therefore, as $n \to \infty$,

$$
P(X_n \leq x) = \sum_{k=0}^{n-1} \frac{x}{n} f(t_k) = x \sum_{k=0}^{n-1} \frac{1}{n} f(t_k) \to x \int_{0}^{1} f(t) \, dt = x,
$$

where convergence of the Riemann sums to the integral holds due to the assumptions on $f$.

Thus the DF of $X_n$ converge to the DF of $U[0, 1]$ at any point $x \in (0, 1)$, and hence everywhere. So $X_n \overset{d}{\to} X \sim U[0, 1]$, whatever the (continuous & bounded) density $f$!

Re the other modes: bad luck, no convergence. This can be seen from the fact that $X_n$ don’t form a Cauchy sequence (in the resp. sense), and so convergence cannot hold. DIY.