1] We have, by definition, for $i=1,\ldots,n$,
\[ p_{X_i\leq t} = \begin{cases} 1 & \text{if } X_i \leq t \\ 0 & \text{if } X_i > t \end{cases} \text{ proba } P(X_i \leq t) = F(t) \]
so they are Bernoulli r.v. with parameter $F(t)$.
They are independent because the $X_i$'s are independent.

2] Apply Hoeffding's inequality, we get, for every $\varepsilon > 0$
\[
P\left( \left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{X_i \leq t} - F(t) \right| \geq \varepsilon \right) \leq 2 \exp\left(-2\varepsilon^2 n\right)
\]
\[
= \mathbb{P}_{n}(t)
\]
Since the right-hand side does not depend on $t$, we may write
\[
\sup_{t \in \mathbb{R}} P\left( \left| \frac{1}{n} F_n(t) - F(t) \right| \geq \varepsilon \right) \leq \sup_{t \in \mathbb{R}} 2 \exp\left(-2\varepsilon^2 n\right)
\]
\[
= 2 \exp\left(-2\varepsilon^2 n\right).
\]

3] We always have
\[
\left| \frac{1}{n} F_n(t) - F(t) \right| \leq \sup_{t \in \mathbb{R}} \left| \frac{1}{n} F_n(t) - F(t) \right|,
\]
so
\[
P\left( \left| \frac{1}{n} F_n(t) - F(t) \right| \geq \varepsilon \right) \leq P\left( \sup_{t \in \mathbb{R}} \left| \frac{1}{n} F_n(t) - F(t) \right| \geq \varepsilon \right).
\]
We can take the supremum on $t$ in the left-hand side, the right-hand side does not depend on $t$! (I know, it's surprising, but observe that for any quantity $A$, $\sup_{t} A(t)$ can also be written $\sup_{s} A(s)$ or whatever letter you want.)
So
\[
\sup_{t \in \mathbb{R}} P\left( \left| \frac{1}{n} F_n(t) - F(t) \right| \geq \varepsilon \right) \leq P\left( \sup_{t \in \mathbb{R}} \left| \frac{1}{n} F_n(t) - F(t) \right| \geq \varepsilon \right),
\]
and thus inequality (3) implies inequality (e).
4] No and no. Proof:

- We can compute

\[
\mathbb{E}[\hat{F}_n(t)] = \frac{1}{\varepsilon} \left( \mathbb{E}[\hat{F}_n(t+\varepsilon)] - \mathbb{E}[\hat{F}_n(t)] \right)
\]

\[
= \frac{1}{\varepsilon} \left( F(t+\varepsilon) - F(t) \right)
\]

Because \( \hat{F}_n(\varepsilon) \) is an unbiased estimator of \( F(\varepsilon) \) for all \( \varepsilon \).

This is "close to" \( F(t) \) for \( \varepsilon \) small, but not equal to it (in general).

So, \( \lim_{n \to \infty} \mathbb{E}[\hat{P}_n(t)] \neq F'(t) = \rho(t) \). Asymptotically biased.

- We know \( \hat{F}_n(a) \) is a consistent estimator of \( F(a) \), so

\[
\hat{F}_n(t+\varepsilon) \xrightarrow{\text{IP}} \frac{1}{\varepsilon} \left( F(t+\varepsilon) - F(t) \right)
\]

and thus

\[
\hat{P}_n(t) \xrightarrow{\text{IP}} \frac{1}{\varepsilon} \left( F(t+\varepsilon) - F(t) \right)
\]

For the same reason, this is not \( = F'(t) \). (So \( \hat{P}_n(t) \) is not consistent.)

5] We can still write, for any \( n, \varepsilon_n \)

\[
\mathbb{E}[\hat{F}_n(t+\varepsilon_n)] = F(t+\varepsilon_n)
\]

Because \( \hat{F}_n \) is unbiased.

We also have \( \mathbb{E}[\hat{F}_n(t)] = F(t) \).

So for any fixed \( t \), for any \( n \), we have

\[
\mathbb{E}[\hat{P}_n(t)] = \frac{1}{\varepsilon_n} \left( F(t+\varepsilon_n) - F(t) \right)
\]

Then, sending \( n \to \infty \), since \( \varepsilon_n \to 0 \), we have

\[
\lim_{n \to \infty} \mathbb{E}[\hat{P}_n(t)] = \lim_{n \to \infty} \left( \frac{1}{\varepsilon_n} \left( F(t+\varepsilon_n) - F(t) \right) \right) = F'(t) = \rho(t)
\]

So \( \hat{P}_n(t) \) is asymptotically unbiased.
Suppose we are sampling the uniform distribution on \([0, 1]\). Typically we will get \(n\) data points that are (approximately) uniformly spread on \([0, 1]\), so \(\hat{F}_n(t)\) looks like

\[
\frac{1}{n} \uparrow \quad \frac{1}{k} \quad \frac{1}{n} \quad \frac{2}{k} \quad \frac{2}{n} \quad \cdots \quad \frac{k}{n} \quad \frac{k}{n} \quad \cdots \quad \frac{n-1}{n} \quad 1
\]

\(0 \quad 1/n \quad 2/n \quad 3/n \quad \cdots \quad n-1/n \quad n\)

Data points \(\sim \frac{k}{n}\) for \(k = 1, \ldots, n\).

For \(t\) arbitrary in \([0, 1]\), \(\hat{F}_n(t + n^{-100})\) and \(\hat{F}_n(t)\) are equal, except if \(t\) is near a data point (more precisely, except if \(t\) is at distance \(\leq n^{-100}\) on the left of a data point).

\(\hat{F}_n\) jumps when we reach a data point

If \(t\) is in this zone,

\[
\hat{F}_n(t + n^{-100}) - \hat{F}_n(t) = \text{jump} = \frac{1}{n}
\]

So "most of the time", \(\frac{\hat{F}_n(t + n^{-100}) - \hat{F}_n(t)}{n^{-100}} = 0\).

At some "exceptional" times,

\[
\frac{\hat{F}_n(t + n^{-100}) - \hat{F}_n(t)}{n^{-100}} = \frac{1}{n} \underbrace{n^{-100}}_{\text{very large}} \rightarrow \frac{89}{n}
\]

In expectation, things work because

Often \(\rightarrow 0\) + Rarely \(\rightarrow \infty\) + Very large \(\rightarrow \frac{89}{n}\).

However, "in probability", we only see the "often" part, and choosing \(\varepsilon_n\) too small leads to an estimation \(\hat{F}_n \frac{1}{n + \varepsilon_n} \rightarrow 0\). Not consistent!
We want to prove that \( \hat{P}_n(t) \) is a consistent estimator of \( P(t) \), which means \( \hat{P}_n(t) \xrightarrow{n \to \infty} P(t) \).

So we fix \( \delta > 0 \), and we want to prove

\[
P \left( \left| \hat{P}_n(t) - P(t) \right| > \delta \right) \xrightarrow{n \to \infty} 0.
\]

Let us write the definition of \( \hat{P}_n(t) \):

\[
\hat{P}_n(t) = \frac{\hat{F}_n(t+n^{-1/4}) - \hat{F}_n(t)}{n^{-1/4}}.
\]

We also have, since \( P(t) = \frac{F(t+n^{-1/4}) - F(t)}{n^{-1/4}} + h(n) \), with \( h(n) \xrightarrow{n \to \infty} 0 \),

by definition of a derivative.

So

\[
\left| \hat{P}_n(t) - P(t) \right| \leq \frac{|\hat{F}_n(t+n^{-1/4}) - F(t+n^{-1/4})|}{n^{-1/4}} + \frac{|\hat{F}_n(t) - F(t)|}{n^{-1/4}} + |h(n)|
\]

by triangular inequality.

By inequality (2), we have, choosing \( \epsilon = \frac{\delta}{10} n^{-1/4} \):

\[
P \left( \left| \hat{F}_n(t) - F(t) \right| \geq \frac{\delta}{10} n^{-1/4} \right) \xrightarrow{n \to \infty} 0
\]

\[
\leq 2 \exp \left( -2 \frac{\delta^2}{100} n^{-1/2}. n \right) = 2 \exp \left( -\frac{\delta^2}{50} n^{1/2} \right)
\]

Similarly, we have

\[
P \left( \left| \hat{F}_n(t+n^{-1/4}) - F(t+n^{-1/4}) \right| \geq \frac{\delta}{10} n^{-1/4} \right) \leq 2 \exp \left( -\frac{\delta^2}{50} n^{1/2} \right).
\]

So

\[
P \left( \left| \frac{\hat{F}_n(t+n^{-1/4}) - F(t+n^{-1/4})}{n^{-1/4}} + \frac{|\hat{F}_n(t) - F(t)|}{n^{-1/4}} \geq \frac{\delta}{5} \right) \xrightarrow{\text{Union bound}} 2 \cdot \exp \left( -\frac{\delta^2}{50} n^{1/2} \right).
\]

Since \( h(n) \to 0 \), it is less than \( \frac{\delta}{5} \) for \( n \) large enough.
In conclusion, 
\[
\frac{1}{n^{1/4}} \left| F_n(t + n^{-1/4}) - F(t + n^{-1/4}) \right| + \frac{1}{n^{-1/4}} \left| F_n(t) - F(t) \right| + h(n)
\]

is less than \( \frac{\delta}{5} + \frac{\delta}{5} = \frac{2\delta}{5} \leq \delta \), for \( n \) large enough, with probability \( \geq 1 - 4 \exp\left(\frac{-\delta^2}{50} n^{1/2}\right) \) \( \rightarrow 0 \) as \( n \to \infty \).

and thus \( \Pr\left( |F_n(t) - F(t)| > \delta \right) \rightarrow 0 \) as \( n \to \infty \).