Theorem (Glivenko-Cantelli)**.** *Let F_n be the empirical distribution function of a sample of size n, for $(X_i)_{i\geq 1} \in \mathbf{R}$ i.i.d random variables with distribution function F. Then*

$$\lim_{n\to\infty} \sup_{x\in\mathbf{R}} |F_n(x) - F(x)| = 0 \quad a.s.$$

Proof. Let $t_1, \dots, t_k \in \mathbf{R}$. Then since $\forall j \leq k \quad (\mathbf{1}_{(-\infty,t_j]}(X_i))_{i\geq 1}$ are also i.i.d random variables. By the strong law of large numbers we know that

$$\forall j \leq k: \quad |F_n(t_j) - F(t_j)| \to 0 \quad (a.s.). \tag{1}$$

So for every $j \leq k$ we have a set $A_j \in \Omega$ with $\mathbf{P}(A_j) = 1$ such that (1) holds. Clearly $\mathbf{P}(\bigcap_{j\leq k} A_j) = 1$, so (by choosing the largest N_j in the definition of the limits) we have:

$$\max_{j=1,\dots,k} |F_n(t_j) - F(t_j)| \to 0 \quad (a.s.). \tag{2}$$

Now, let $h \nearrow t_j$. Then

$$\begin{aligned} F_n(t_j^-) &= \lim_{h\nearrow t_j} \frac{1}{n} \sum_{i=1,\dots,n} \mathbf{1}_{(-\infty,h]}(X_i) \\ &= \frac{1}{n} \sum_{i=1,\dots,n} \mathbf{1}_{(-\infty,t_j)}(X_i) \end{aligned}$$

Because $(\mathbf{1}_{(-\infty,t_j)}(X_i))_{i\geq 1}$ are i.i.d random variables (with finite expectation), the strong law of large numbers gives us:

$$|F_n(t_j^-) - F(t_j^-))| = |F_n(t_j^-) - \mathbf{P}(X_m \in (-\infty, t_j))| \to 0 \quad (a.s.).$$

Using the same argument as before we can now conclude:

$$\max_{j=1,\dots,k} |F_n(t_j^-) - F(t_j^-)| \to 0 \quad (a.s.). \tag{3}$$

Continuing, fix any $\varepsilon > 0$ and choose $t_j = \inf\{t \in \mathbf{R} : F(t) \geq j\varepsilon\}$ for $i = 1, \dots, \lfloor \frac{1}{\varepsilon} \rfloor$. (Note that $t_0 = -\infty$). Then $\forall t \in \mathbf{R}$ there is a $j \in \mathbf{N}$ with $t \in (t_{j-1}, t_j)$ since F is a cdf and $j\varepsilon \leq 1$. Now we estimate using $t < t_i$ and $t > t_{i-1}$:

$$\begin{aligned} F_n(t) - F(t) &\leq F_n(t_j^-) - F(t_{j-1}) \\ &\leq (F_n(t_j^-) - F(t_j^-)) + (F(t_j^-) - F_n(t_{j-1})) + (F_n(t_{j-1}) - F(t_{j-1})) \\ &\leq \max_{j=1,\dots,k} |F_n(t_j^-) - F(t_j^-)| + (F(t_j^-) - F_n(t_{j-1})) + \max_{j=1,\dots,k} |F_n(t_j) - F(t_j)| \\ &\leq \max_{j=1,\dots,k} |F_n(t_j^-) - F(t_j^-)| + j\varepsilon - F_n(t_{j-1}) + \max_{j=1,\dots,k} |F_n(t_j) - F(t_j)| \end{aligned}$$

By (1) we have for all $n \geq N(\varepsilon)$, that $F(t_{j-1}) - F_n(t_{j-1}) \leq \varepsilon \iff -F_n(t_{j-1}) \leq \varepsilon - F(t_{j-1}) \leq \varepsilon - (j-1)\varepsilon$, almost surely. We conclude, that

$$F_n(t) - F(t) \leq \max_{j=1,\dots,k} |F_n(t_j^-) - F(t_j^-)| + 2\varepsilon + \max_{j=1,\dots,k} |F_n(t_j) - F(t_j)| \quad (a.s.).$$

A completely symmetrical argument then shows:

$$|F_n(t) - F(t)| \leq \max_{j=1,\dots,k} |F_n(t_j^-) - F(t_j^-)| + 2\varepsilon + \max_{j=1,\dots,k} |F_n(t_j) - F(t_j)| \quad (a.s.). \tag{4}$$

Note that the choice of the numbers t_j only depends on ε. Therefore for $\varepsilon > 0$ (4) implies

$$\sup_{t\in\mathbf{R}} |F_n(t) - F(t)| \leq \max_{j=1,\dots,k} |F_n(t_j^-) - F(t_j^-)| + 2\varepsilon + \max_{j=1,\dots,k} |F_n(t_j) - F(t_j)| \quad (a.s.).$$

For N large enough. Combining this with (2) and (3) and choosing $N_1 \geq N$ large enough (also greater than the N necessary in (2) and (3)), yields:

$$\sup_{t\in\mathbf{R}} |F_n(t) - F(t)| \leq \varepsilon + 2\varepsilon + \varepsilon \text{ for } n \geq N_1 \quad (a.s.).$$

Therefore we also have

$$\limsup_{n\to\infty} \sup_{t\in\mathbf{R}} |F_n(t) - F(t)| \leq +5\varepsilon \quad (a.s.).$$

Because $\varepsilon > 0$ was arbitrary, we conclude

$$0 \leq \liminf_{n\to\infty} \sup_{t\in\mathbf{R}} |F_n(t) - F(t)| \leq \limsup_{n\to\infty} \sup_{t\in\mathbf{R}} |F_n(t) - F(t)| = 0 \quad (a.s.),$$

which yields the claim. □