Connections between Convergence in MSE and AE

Xintong Li

April 14, 2025

This section establishes the connections between the convergence in MSE and the convergence in probabilistic order for AE in the context of nested simulation procedures. In order to show the connections between the convergence in MSE and the convergence in probabilistic order for AE, we first need to state the definition for a sequence of random variables to converge in those two forms.

Definition 1 *Let $\hat{\rho}_\Gamma$ be an estimator of ρ with a simulation budget of Γ. We write $\mathbb{E}\left[(\hat{\rho}_\Gamma - \rho)^2\right] = \mathcal{O}\left(\Gamma^{-\xi}\right)$, that is, $\hat{\rho}_\Gamma$ converges in MSE to ρ in order ξ if there exists a constant C such that*

$$\limsup_{\Gamma\to\infty} \frac{\mathbb{E}\left[(\hat{\rho}_\Gamma - \rho)^2\right]}{\Gamma^{-\xi}} \leq C.$$

Definition 2 *Let $\hat{\rho}_\Gamma$ be an estimator of ρ with a simulation budget of Γ. We write $|\hat{\rho}_\Gamma - \rho| = \mathcal{O}_{\mathbb{P}}(\Gamma^{-\xi})$, that is $\hat{\rho}_\Gamma$ converges in probabilistic order ξ to ρ if for a sufficiently large Γ, for all $\epsilon > 0$ there exists a constant C such that*

$$\mathbb{P}\left(|\hat{\rho}_\Gamma - \rho| \geq C\Gamma^{-\xi}\right) \leq \epsilon.$$

We start our analysis by showing the convergence in probabilistic order from the convergence in MSE. Let $\hat{\rho}_\Gamma$ be an estimator of ρ with a simulation budget of Γ, and assume that $\mathbb{E}\left[(\hat{\rho}_\Gamma - \rho)^2\right] = \mathcal{O}\left(\Gamma^{-\xi}\right)$. Then, from the definition of convergence in MSE, there exists a constant C such that

$$\limsup_{\Gamma} \frac{\mathbb{E}\left[(\hat{\rho}_\Gamma - \rho)^2\right]}{\Gamma^{-\xi}} \leq C.$$

Hence, there exists some Γ such that for all $\gamma \geq \Gamma$,

$$\mathbb{E}\left[(\hat{\rho}_\gamma - \rho)^2\right] \leq C\gamma^{-\xi}.$$

The convergence in probabilistic order can be shown by separating the expectation into two parts: tail and non-tail components.

$$\mathbb{E}\left[(\hat{\rho}_\gamma - \rho)^2 \cdot \mathbb{I}_{\{|\hat{\rho}_\gamma - \rho| \leq d\gamma^s\}}\right] + \mathbb{E}\left[(\hat{\rho}_\gamma - \rho)^2 \cdot \mathbb{I}_{\{|\hat{\rho}_\gamma - \rho| > d\gamma^s\}}\right] \leq C\gamma^{-\xi}.$$

The first term is always positive, and the second term can be bounded from below by the indicator function.

$$\begin{aligned}
\mathbb{E}\left[(\hat{\rho}_\gamma - \rho)^2 \cdot \mathbb{I}_{\{|\hat{\rho}_\gamma - \rho| > d\gamma^s\}}\right] &\geq \mathbb{E}\left[d^2\gamma^{2s} \cdot \mathbb{I}\{|\hat{\rho}_\gamma - \rho| > d\gamma^s\}\right] \\
&= d^2\gamma^{2s} \cdot \mathbb{E}\left[\mathbb{I}\{|\hat{\rho}_\gamma - \rho| > d\gamma^s\}\right] \\
&= d^2\gamma^{2s} \cdot \mathbb{P}\left(|\hat{\rho}_\gamma - \rho| > d\gamma^s\right).
\end{aligned}$$

Combining bounds on the two terms, we have

$$d^2\gamma^{2s}\mathbb{P}\left(|\hat{\rho}_\gamma - \rho| > d\gamma^s\right) \leq C\gamma^{-\xi}.$$

Let $s = -\frac{\xi}{2}$. Arranging the terms, we have

$$\mathbb{P}\left(|\hat{\rho}_\gamma - \rho| > d\gamma^{-\frac{\xi}{2}}\right) \leq \frac{C}{d^2}.$$

Hence, for all $\epsilon > 0$, there exist $C^* = \sqrt{\frac{C}{\epsilon}}$ such that for all $\gamma \geq \Gamma$,

$$\mathbb{P}\left(|\hat{\rho}_\gamma - \rho| > C^* \gamma^{-\frac{\xi}{2}}\right) \leq \epsilon$$

In essence, the above expression is a definition of convergence in probabilistic order, that is,

$$|\hat{\rho}_\gamma - \rho| = \mathcal{O}_{\mathbb{P}}\left(\Gamma^{-\frac{\xi}{2}}\right).$$

Theorem 1 *Let $\hat{\rho}_\Gamma$ be an estimator of ρ with a simulation budget of Γ. If $\hat{\rho}_\Gamma$ converges in MSE to ρ in order ξ, then $\hat{\rho}_\Gamma$ converges in probabilistic order to ρ in order $\frac{\xi}{2}$.*

To the best of our knowledge, Theorem 1 has not been explicitly stated in the literature. It is the first result that shows the connections between the convergence in MSE and the convergence in probabilistic order for AE in the context of nested simulation. Theorem 1 is a general result that can be applied to any nested simulation procedure that converges in MSE to ρ in order ξ. If the estimator converges in MSE to ρ in order ξ, then it converges in probabilistic order to ρ in order $\frac{\xi}{2}$.

While the convergence in MSE implies the convergence in probabilistic order, the converse is not necessarily true. Similarly, the above argument is applied in reverse. Let $\hat{\rho}_\Gamma$ be an estimator of ρ with a simulation budget of Γ, and assume that $|\hat{\rho}_\Gamma - \rho| = \mathcal{O}_{\mathbb{P}}(\Gamma^{-\xi})$. The MSE of $\hat{\rho}_\Gamma$ can be separated into the same two parts.

$$\mathbb{E}\left[(\hat{\rho}_\Gamma - \rho)^2\right] = \mathbb{E}\left[(\hat{\rho}_\Gamma - \rho)^2 \cdot \mathbb{I}_{\{|\hat{\rho}_\Gamma - \rho| \leq d\Gamma^{-2\xi}\}}\right] + \mathbb{E}\left[(\hat{\rho}_\Gamma - \rho)^2 \cdot \mathbb{I}_{\{|\hat{\rho}_\Gamma - \rho| > d\Gamma^{-2\xi}\}}\right],$$

where the first term can be bounded from above.

$$\begin{aligned}\mathbb{E}\left[(\hat{\rho}_\Gamma - \rho)^2 \cdot \mathbb{I}_{\{|\hat{\rho}_\Gamma - \rho| > d\Gamma^s\}}\right] &\leq d^2\Gamma^{-2\xi} \cdot \mathbb{E}\left[\mathbb{I}\{|\hat{\rho}_\Gamma - \rho| > d\Gamma^s\}\right] \\ &= d^2\Gamma^{-2\xi} \cdot \mathbb{P}\left(|\hat{\rho}_\Gamma - \rho| > d\Gamma^s\right) \leq d^2\Gamma^{-2\xi}.\end{aligned}$$

However, the second term is not always bounded. If the random variable $\hat{\rho}_\Gamma$ admits a density function f, then the second term can be further decomposed as follows.

$$\mathbb{E}\left[(\hat{\rho}_\Gamma - \rho)^2 \cdot \mathbb{I}_{\{|\hat{\rho}_\Gamma - \rho| > d\Gamma^{-2\xi}\}}\right] = \int_{-\infty}^{-d\Gamma^{-2k}} (x-\rho)^2 f(x)dx + \int_{d\Gamma^{-2\xi}}^{\infty} (x-\rho)^2 f(x)dx.$$

Hence, $\hat{\rho}_\Gamma$ converges in MSE to ρ in order 2ξ if and only if both integrals converge in order higher than 2ξ. The above argument shows that the convergence in probabilistic order does not necessarily imply the convergence in MSE. Hence, the converse of Theorem 1 is not necessarily true, and the convergence in probabilistic order is a weaker form of convergence than the convergence in MSE.