基礎数理学統計レポート

1114230018

原田雄大

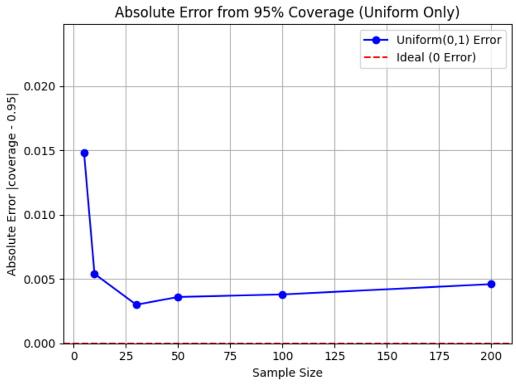
概要

本シミュレーションは **Python（NumPy, SciPy, pandas, matplotlib）** を用いて実施した。
１と２ではそれぞれのサンプルサイズ：(Size[5,10,30,50,100,200]）について，区間推定結果が正しい確率（Coverage Rate）の関係を求めた．また，95%からの誤差の絶対値をグラフにした。３ではそれぞれ Coverage Rate の 95%からの誤差(Uniform Error)と(Exponential Error)との結果を比較して考察している．

本文

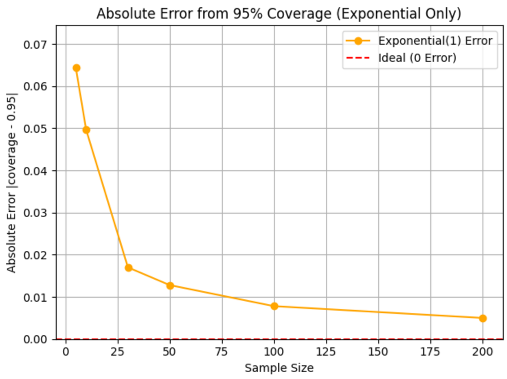
１．Uniform：$p.d.f.\ f(x) = \begin{cases} 1 & (0 \le x \le 1) \\ 0 & (otherwise) \end{cases}$

Size	Coverage Rate (Uniform)
5	0.9532
10	0.9446
30	0.9470
50	0.9464
100	0.9462
200	0.9546



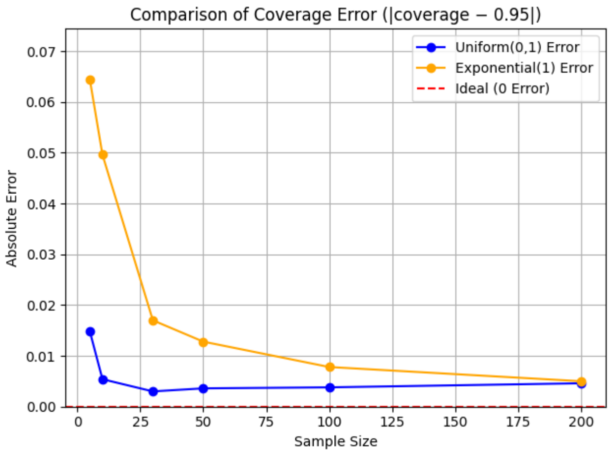
２．Exponential：$p.d.f.\ f(x) = \begin{cases} e^{-x} & (x \ge 0) \\ 0 & (otherwise) \end{cases}$

Size	Coverage Rate (Exponential)
5	0.8856
10	0.9004
30	0.9330
50	0.9372
100	0.9422
200	0.9450



３．Uniform と Exponential の比較

Size	Uniform Error	Exponential Error
5	0.0148	0.0644
10	0.0054	0.0496
30	0.0030	0.0170
50	0.0036	0.0128
100	0.0038	0.0078
200	0.0046	0.0050



本シミュレーションでは、t 分布を用いた 95%信頼区間のカバレッジ率を一様分布と指数分布で比較した。一様分布は分布が対称かつ有界であるため、標本平均が早期に正規分布に近づきやすく、全体的にカバレッジ率は 0.95 をわずかに上回る結果となった。これは t 分布による区間推定がやや保守的で、真の平均を含みやすいことを示している。一方、指数分布では分布が非対称であるため、特に n が小さい場合にカバレッジ率が大きく理論値を下回った。これは t 分布の前提である正規性が崩れているためである。ただし、n が増加すると中心極限定理の影響により両分布とも標本平均が正規に近づき、n=200 ではカバレッジ率が一致した。以上より、t 分布を用いた区間推定の精度は分布形状とサンプルサイズに強く依存することが確認された。

本シミュレーションで用いたコードについては
GitHub（https://github.com/yu531deve/FundamentalMathematicalStatistics）
で公開している。