CLT and the Pareto Distribution: Deep Dive

CLT and Pareto — the critical blurb

Classical CLT: If X■, X■, … are i.i.d. with mean μ and finite variance σ^2, then: Z■ = $\sqrt{}$n (X■■ – μ)/σ $\rightarrow$ ■(0,1) in distribution.

Pareto heavy tails: Pareto(α, x■) has tail P(X > x) ~ (x■/x)^α. - Mean finite iff $\alpha > 1$, μ = αx■/(α–1). - Variance finite iff $\alpha > 2$, σ^2 = αx■² / ((α–1)² (α–2)).

Failure mode: For $1 < \alpha \leq 2$, the mean exists but variance is infinite $\rightarrow$ classical CLT fails. The generalized CLT applies: properly scaled sums converge to an α-stable law (non-Gaussian). Studentizing by sample sd does not restore normality: t-statistics remain heavy-tailed.

Practical consequences: - t-intervals for the mean under-cover severely when $\alpha \leq 2$. - Robust estimators (trimmed means, Winsorized means, Huber M-estimators) regain asymptotic normality.

Takeaways for demonstrations: 1. $\alpha = 2.5$ (finite variance): LLN + CLT work as expected $\rightarrow$ sample mean normalized looks Gaussian. 2. $\alpha = 1.5$ (finite mean, infinite variance): LLN holds but CLT fails $\rightarrow$ histograms remain heavy-tailed, Q–Q vs Normal bends, coverage of 95% CIs is poor. 3. Trimmed means dramatically stabilize inference in heavy tails.

This contrast makes Pareto the textbook demonstration of where LLN still works but CLT fails spectacularly.