Heuristics: Where does normality of ε come from?

- Poincare: "Everyone believes in the *Gaussian* law of errors, the experimentalists because they think it is a mathematical theorem, the mathematicians because they think it is an experimental fact."

- Gauss (1809) worked backwards to construct a distribution of errors for which the least squares is the maximum likelihood estimate ("the most probable estimate"). Hence normal distribution is sometimes called Gaussian.

- Central limit theorem justification: De-Moivre-Laplace, Liapunov, Levy, Khinchin, and Lindeberg (no Gauss here). In econometrics, Haavelmo, in his "Probability Approach to

Econometrics", *Econometrica* 1944, was a prominent proponent of this justification.

- Under the CLT justification, the errors ε_i are thought of as a sum of a large number of small and independent elementary errors v_j, and therefore will be approximately Gaussian due to the central limit theorem considerations.

- If elementary errors $v_j, j = 1, 2, ...$ are i.i.d. mean-zero and $E[v_j^2] < \infty$, then for *large* N

$$\varepsilon_i = \sqrt{N} \left[\frac{\sum_{j=1}^N v_j}{N} \right] \approx_d N(0, Ev_j^2),$$

  as follows from the CLT.

- However, if elementary errors v_j are i.i.d. symmetric and $E[v_j^2] = \infty$, then *for large*

n (with additional technical restrictions on the tail behavior of v_j).

$$\varepsilon_i = N^{1-\frac{1}{\alpha}} \left[\frac{\sum_{j=1}^{N} v_j}{N} \right] \approx_d Stable$$

where α is the largest finite moment: $\alpha = \sup\{p : E|v_j|^p < \infty\}$. This follows from the CLT proved by Khinchine and Levy. The Stable distributions are also called sum-stable and *Pareto-Levy* distributions.

- Densities of symmetric stable distributions are "bell-shaped" but have thick tails which behave approximately like power functions $x \mapsto$ const $\cdot |x|^{-\alpha}$ in the tails, with $\alpha < 2$.

- Another interesting side observation: If $\alpha > 1$, the sample mean $\sum_{j=1}^{N} v_j/N$ is a converging statistic, if $\alpha < 1$ the sample mean $\sum_{j=1}^{N} v_j/N$ is a diverging statistics.

- References: Embrechts et al. *Modelling Extremal Events*