Financial Econometrics

Professor Martins

University of Colorado at Boulder

Fall 2023

Lecture 7

Testing for Normality

- If $r = \mu + \sigma Z$ with $\sigma > 0$ and $Z \sim N(0,1)$ then $\alpha \in (0,1)$

$$F_r^{-1}(\alpha) = \mu + \sigma F_Z^{-1}(\alpha), \text{ see slides for lecture 3}$$

- Consider a random sample $\{r_t\}_{t=1}^n$ and order these observations from the smallest to the largest. We denote the smallest to be $r_{(1)}$, the second smallest to be $r_{(2)}$ until the largest $r_{(n)}$. Clearly,

$$r_{(1)} \leq r_{(2)} \leq \cdots \leq r_{(n)}.$$

  The set $\{r_{(t)}\}_{t=1}^n$ is called the "order statistics" associated with $\{r_t\}_{t=1}^n$.

Testing for Normality

- If $F_n(r) = \frac{1}{n}\sum_{t=1}^{n} I_{\{r_t \leq r\}}$, we know that $F_n(r) \xrightarrow{p} F_r(r)$.
- It is easy to show that the quantile $q_{F_n}(\alpha)$ of F_n is

$$q_{F_n}(\alpha) = \begin{cases} r_{(n)} & \text{if } \frac{n-1}{n} < \alpha < 1 \\ r_{(n-1)} & \text{if } \frac{n-2}{n} < \alpha \leq \frac{n-1}{n} \\ \vdots & \vdots \\ r_{(2)} & \text{if } \frac{1}{n} < \alpha \leq \frac{2}{n} \\ r_{(1)} & \text{if } 0 < \alpha \leq \frac{1}{n} \end{cases}$$

  Hence,

$$F_r^{-1}(\alpha) \approx q_{F_n}(\alpha) \text{ and } q_{F_n}(\alpha) \approx \mu + \sigma F_Z^{-1}(\alpha)$$

- If r is normally distributed, any difference between $q_{F_n}(\alpha)$ and $\mu + \sigma F_Z^{-1}(\alpha)$ is due to the fact that $F_r^{-1}(\alpha) \neq q_{F_n}(\alpha)$

QQ-plots

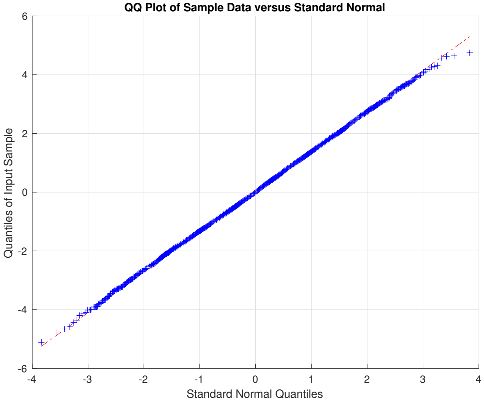

Figure: QQ plot normal

QQ-plots

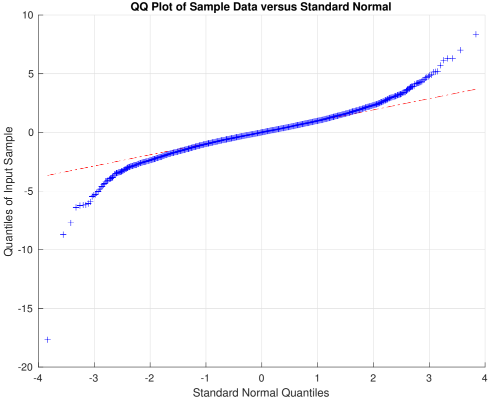

Figure: QQ plot Student-t, E(T)=0, scale= $\sqrt{5/8}$

QQ-plots

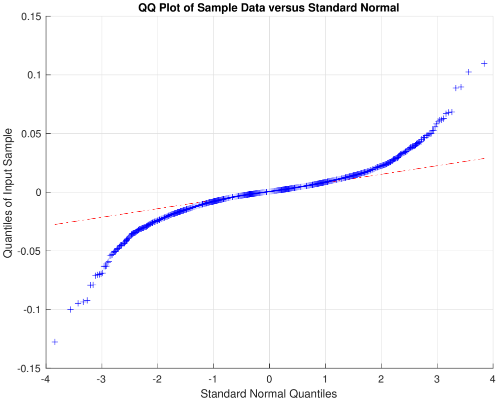

Figure: QQ plot S&P 500 log-returns

Kolmogorov-Smirnoff Test

We know that for fixed r,

$$F_n(r) - F_r(r) \xrightarrow{p} 0.$$

In fact, a fundamental result in Statistics (Glivenko-Cantelli Theorem), says that

$$\sup_{r \in \mathbb{R}} |F_n(r) - F_r(r)| \xrightarrow{p} 0.$$

Kolmogorov showed that for $0 < z < \infty$ and any F continuous

$$P\left(\sqrt{n} \sup_{r \in \mathbb{R}} |F_n(r) - F_r(r)| \leq z\right) = P\left(\sqrt{n} \sup_{y \in [0,1]} |\frac{1}{n} \sum_{t=1}^{n} I_{\{U_t \leq y\}} - y| \leq z\right)$$

and that

$$P\left(\sqrt{n} \sup_{r \in \mathbb{R}} |F_n(r) - F_r(r)| \leq z\right) \to H(z) = 1 - 2 \sum_{i=1}^{\infty} (-1)^{i-1} \exp(-2i^2 z^2)$$

Kolmogorov-Smirnoff Test

Let

$$H_0 : F_r = F_0 \text{ and } H_1 : F_r \neq F_0$$

where F_0 is the normal cumulative distribution with parameters μ_0, σ_0^2 and

$$D_n = \sqrt{n} \sup_{r \in \mathbb{R}} |F_n(r) - F_r(r)|$$

is the test statistic. If the null hypothesis is true, then the distribution of D_n is approximately $H(z)$. Hence, if $D_n \leq c$ we accept H_0 and if $D_n > c$ we reject H_0. c depends on the test level of significance, and can be found by

$$P(D_n \geq c_\alpha | H_0) = \alpha \text{ for } \alpha \in (0, 1)$$

Implementation using MATLAB

- Data: S&P500 index value from January, 1988 to June, 2020
- Code: MLESP500_KS.m
- Key functions in MATLAB: makedist, kstest