Financial Econometrics

Professor Martins

University of Colorado at Boulder

Fall 2023

Lecture 2

Stochasticity of returns

- Since at time $t - h$ it is not possible to know the price at time t, there is uncertainty about $R_{t,h}$, $r_{t,h}$ and $\rho_{t,h}$.
- Particularly problematic is the event $P_t < P_{t-h}$, implying

$$R_{t,h},\ r_{t,h} < 0 \text{ or } \rho_{t,h} < 1.$$

- The risk of losing money from buying an asset in period $t - h$ for P_{t-h} and selling it in period t for $P_t < P_{t-h}$ is the fundamental problem of investing.
- We will think of $R_{t,h}$, $r_{t,h}$ and $\rho_{t,h}$ as random variables.

Stochasticity of returns

- The insight that returns can be viewed as random variables is the foundation of our approach to the study of empirical finance (see Bachelier (1900)).
- It allows for the construction of statistical models that represent the evolution of returns through time and for the testing and evaluation of various models developed in theoretical Finance.

Example

At any time period t a collection (sample) of n observed log-returns on a financial asset can be represented by

$$\{r_{t-(n-1)}, \cdots, r_{t-1}, r_t\}$$

Without loss of generality, set $t = n$ and write

$$\{r_1, \cdots, r_{n-1}, r_n\} = \{r_t\}_{t=1}^{n}.$$

The following figure shows a graph of one such sample associated with log-returns on the S&P 500 (see MATLAB code `S&P_log_returns.m` that produces this figure).

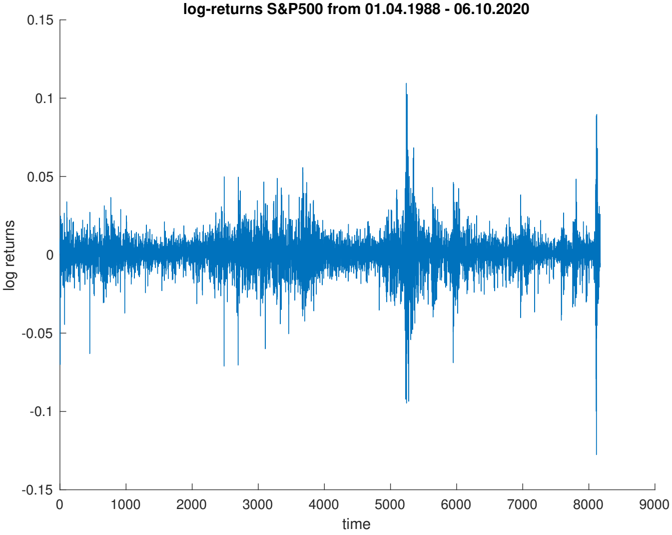

Figure: Graph of log returns on S&P 500

Stochasticity of returns

Suppose

- $\{r_t\}_{t=1}^{n}$ is a collection of independently and identically distributed random variables
- If r_t has a distribution function F, denoted by $r_t \sim F$, then using a sample of past returns, the distribution F can be estimated.

Important practical questions, such as:

- What is $E(r_t)$?
- What is the volatility, i.e., $V(r_t)$?
- What is the probability that $r_t < -a$, i.e., $P(r_t < -a)$ for $a > 0$?

can be answered.

Stochasticity of returns

Satisfactorily answering these questions depends, in turn, on constructing statistical models that reasonably represent the stochastic process that generates prices or returns.

This representation leads to other questions such as:

- Is it reasonable to assume that $\{r_t\}_{t=1}^{n}$ is a collection of independently and identically distributed random variables?
- If $V(r_t)$ exists, can it be suitably expressed as a parametric function of past returns?
- If r_t and r_{t-h} are statistically correlated, is there a parametric model that suitably captures this correlation?

It is apparent that the study of empirical finance depends on a solid understanding of probability and statistics.

Parametrically indexed distributions

- Often we assume that distribution functions belong to a class of functions that are parametrically indexed. In this case, we write

$$F_X(x; \theta_0) \text{ where } \theta_0 \in \Theta \subseteq \mathbb{R}^p.$$

  The set Θ is called the parameter space, θ_0 is called the "true value" of the parameter.

- When X is a continuous random variable we write

$$F_X(x; \theta_0) = \int_{-\infty}^{x} f(u; \theta_0) du$$

  and $f(u; \theta_0)$ is the parametrically indexed density function associated with $F_X(x; \theta_0)$.

Parametrically indexed distributions

- Similarly, when X is a discrete random variable taking values $\{x_1, x_2, \dots\}$ we write the probability function of X as

$$p(x_i; \theta_0) \text{ where } \theta_0 \in \Theta \subseteq \mathbb{R}^p.$$

such that

$$\sum_{i \in \mathbb{N}} p(x_i; \theta_0) = 1 \text{ and } F_X(x; \theta_0) = \sum_{x_i \leq x} p(x_i; \theta_0).$$

Example 1

Let n be the number of trials associated with a stochastic experiment that allows for two outcomes: success with probability θ and failure with probability $1-\theta$. Let X be the total number of success in n trials, then

$$P(X=k) = \left(\begin{array}{c} n \\ k \end{array} \right) \theta^k (1-\theta)^{n-k}$$

for $k = 1, 2, \cdots, n$. $P(X=k)$ is called the binomial distribution with parameters n and θ and is denoted $B(n,\theta)$. We write $X \sim B(n,p)$.

Using the Binomial Theorem, $(x+y)^n = \sum_{j=0}^{n} \left(\begin{array}{c} n \\ j \end{array} \right) x^n y^{n-j}$

we can easily obtain

$$E(X) = n\theta \text{ and } V(X) = n\theta(1-\theta)$$

Example 2

Let X be a continuous random variable taking values in $[a, b] \subset \mathbb{R}$ with $a < b$ and density

$$f(x; a, b) = \frac{1}{b - a}$$

for $x \in [a, b]$ and 0 otherwise. We say that f has support on $[a, b]$. Then,

$$F(x; a, b) = \int_a^x \frac{1}{b-a} du = \frac{x - a}{b - a}$$

and

$$E(X) = \frac{a + b}{2} \text{ and } V(X) = \frac{(b - a)^2}{12}$$

In this case we say that X is uniformly distributed over $[a, b]$ and write $X \sim U[a, b]$

Example 3

- Let X be a continuous random variable that takes values in the interval $(-\infty, \infty)$ with density

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}, \text{ with } \mu \in \mathbb{R} \text{ and } \sigma > 0.$$

  We say that $X \sim N(\mu, \sigma^2)$ with parameters μ and σ^2.
- It can be shown that $E(X) = \mu$ and $V(X) = \sigma^2$. When $\mu = 0$ and $\sigma^2 = 1$ we write $X \sim N(0, 1)$ and say that X has a standard gaussian (normal) density.

Example 4

- If $Z \sim N(0,1)$, then $Y = \mu + \sigma Z$ where $\mu \in \mathbb{R}$ and $\sigma > 0$ is such that $E(Y) = \mu$, $V(Y) = \sigma^2$.
- $f_Y(y) = \frac{1}{\sigma} f_Z\left(\frac{y-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(y-\mu)^2}{\sigma^2}}$.
- $F_Y(y) = \int_{-\infty}^{y} f_Y(\alpha) d\alpha = \int_{-\infty}^{y} \frac{1}{\sigma} f_Z\left(\frac{\alpha-\mu}{\sigma}\right) d\alpha$. Changing variables by letting $z = \frac{\alpha-\mu}{\sigma}$ we have that

$$F_Y(y) = \int_{-\infty}^{y-\mu/\sigma} \frac{1}{\sigma} f_Z(z) \sigma dz = \int_{-\infty}^{y-\mu/\sigma} f_Z(z) dz.$$