Module 9.3: Time Series Analysis
Fall Term 2023

Week 2:

Returns; Autocorrelation; Stationarity

HSLU Lucerne University of Applied Sciences and Arts

Outline in Weeks

1. Introduction; Descriptive Modelling
2. Returns; Autocorrelation; Stationarity
3. ARMA Models
4. Unit Roots; Regressions between Time Series
5. Volatility Modelling
6. Value at Risk
7. Cointegration

Outline

Asset Returns

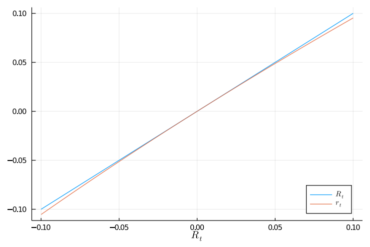
- We consider two definitions of returns:
  1. *Simple* return between dates $t-1$ and t [or: in period t]
  $$R_t = \frac{P_t - P_{t-1}}{P_{t-1}},$$
  where P_t is the asset price at time t.
  2. Continuously compounded return or *log return*
  $$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(1 + R_t).$$
- They are typically very close for daily returns, as
$$r_t = \log(1 + R_t) \approx R_t,$$
when $R_t \approx 0$.

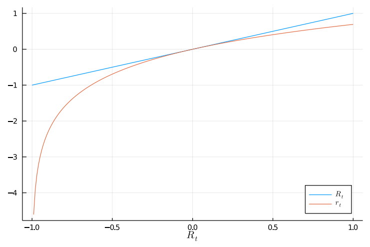
- Log returns and 'simple' returns often are very close, as

$$r_t = \ln(1 + R_t) \approx R_t \text{ when } R_t \approx 0.$$

- Simple returns are bounded below by -1 (100% loss). Log returns live on $(-\infty, \infty)$. Easier to model (e.g., normal distribution).

Simple vs. Log Returns





Log returns: Intuition

- If a one-period interest rate of r is compounded n times, then

$$P_t = (1 + r/n)^n P_{t-1}.$$

- As $n \to \infty$, $(1 + r/n)^n \to e^r$, so

$$P_t = e^r P_{t-1} \Leftrightarrow r = \log(P_t/P_{t-1}) = \ln P_t - \log P_{t-1}.$$

Portfolio Returns

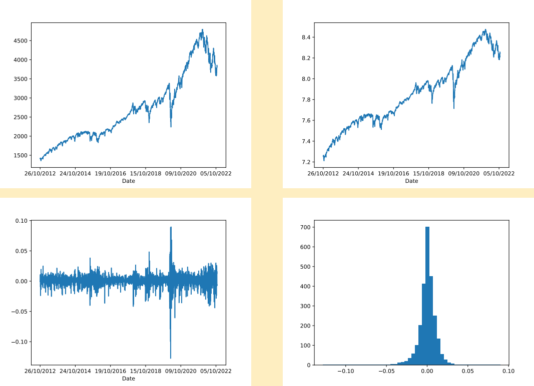
- Advantage of continuously compounded returns: *multi-period return* is sum of single-period returns.
- Advantage of simple returns: *portfolio return* is weighted sum of asset returns.
- Proof: If an investor buys n_i shares in stock i, then the value of the portfolio at time $t-1$ is $V_{t-1} = \sum_{i=1}^{n} n_i P_{i,t-1}$.
- Ignoring dividends, the payoff is $V_t = \sum_{i=1}^{n} n_i P_{i,t}$, so the return on the portfolio is

$$
\begin{aligned}
R_{p,t} = \frac{V_t - V_{t-1}}{V_{t-1}} &= \frac{\sum_{i=1}^{n} n_i (P_{i,t} - P_{i,t-1})}{V_{t-1}} \\
&= \sum_{i=1}^{n} \underbrace{\frac{n_i P_{i,t-1}}{V_{t-1}}}_{w_i} \underbrace{\frac{(P_{i,t} - P_{i,t-1})}{P_{i,t-1}}}_{R_{i,t}} = \sum_{i=1}^{n} w_i R_{i,t}.
\end{aligned}
$$

Stylized Facts of Asset Returns

- Prices display (time-varying) *trend*, and variation proportional to price level (motivation for taking logs).
- Returns have constant mean close to zero, and very little autocorrelation.
- Returns display *volatility clustering*: alternating periods of high and low variability.
- Returns have non-Gaussian distribution, *fat tails* (excess kurtosis).
- Interest rates display long swings, very slow *mean-reversion*.
- Interest rate changes have similar characteristics as returns.

Example: S&P 500 index values and returns, 10/29/2012–10/27/2022



Testing Normality

- Normality can be tested by examining the *skewness* and *kurtosis*.
- Skewness $\mathrm{SK} = m_3/\sqrt{m_2^3}$ and kurtosis $\mathrm{K} = m_4/m_2^2$, where m_j is the j-th centralized moment[1] $m_j = \mathbb{E}[(r_t - \mathbb{E}[r_t])^j]$.
- A normal distribution has SK=0 and K=3.
- *Jarque-Bera normality test*:

$$\mathrm{JB} = \frac{T}{6}\widehat{\mathrm{SK}}^2 + \frac{T}{24}(\widehat{\mathrm{K}} - 3)^2,$$

where the skewness and kurtosis of r_t can be estimated as

$$\widehat{\mathrm{SK}} = \hat{m}_3/\sqrt{\hat{m}_2^3}, \quad \text{and} \quad \widehat{\mathrm{K}} = \hat{m}_4/\hat{m}_2^2, \quad \text{with } \hat{m}_j = \frac{1}{T}\sum_{t=1}^{T}(r_t - \overline{r})^j.$$

- Under the null hypothesis of normality: $\mathrm{JB} \xrightarrow{d} \chi_2^2$.

[1] I.e., the second centralized moment m^2 is just the variance, otherwise known as σ^2.

Outline

Stochastic Processes

- Time series analysis is concerned with modelling, estimating, analyzing and forecasting returns and other financial and economic variables.
- A *time series* $\{y_t, t = 1, 2, \ldots, T\}$ is a collection of subsequent observations on a particular variable. We view such a time series as a *realization* of a *discrete-time stochastic process* $\{Y_t, t = 1, 2, \ldots\}$, which is a collection of (dependent) random variables.
- The goal is to determine which process $\{Y_t\}$ generated the data.
- The distinction between $\{Y_t\}$ (the process) and $\{y_t\}$ (the realization) will usually not be emphasized.
- We will not consider continuous-time stochastic processes here (e.g. Brownian motion).

White Noise

- An important example of a stationary process is the *white noise* process, which has zero mean[2] and zero autocovariances.

$$
\begin{aligned}
\mathbb{E}[U_t] &= 0, \\
\mathrm{var}(U_t) &= \mathbb{E}\left[U_t^2\right] = \sigma^2, \\
\mathrm{cov}(U_t, U_{t-k}) &= \mathbb{E}[U_t U_{t-k}] = 0, \qquad k = 1, 2, \ldots
\end{aligned}
$$

- The notation U_t emphasizes the similarity to regression errors.
- White noise is *unpredictable*.
- It is the building block for other processes (which may be predictable).

[2] Brooks allows a white noise process to have a non-zero mean. Usually such a process is called an *uncorrelated* process.

Outline

Excursus: The Efficient Market Hypothesis

- The *weak form EMH*[3] posits that past prices and returns cannot predict future returns.
- This implies that no fund manager can consistently outperform the market, at least based on historical prices alone.
- If weak form EMH holds, then returns should be *uncorrelated*. Since the mean return is small for daily data, they should therefore resemble wight noise.
- An important application of time series analysis ist testing whether the EMH holds.
- The most basic way to do this is to test whether the returns have been generated by a white noise process.

[3] Fama (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 25(2), pp. 383–417.

Outline

Autocorrelation Function

- Recall that if a process $\{Y_t\}$ is white noise, then it is uncorrelated at all lags (i.e., Y_t should be uncorrelated with Y_{t-1}, with Y_{t-2}, etc.)
- Formally, its *autocorrelation function (ACF)* τ_s is zero for all s, where the ACF is defined from the *autocovariances* γ_s, as

$$\tau_s = \mathrm{Corr}(Y_t, Y_{t-s}) = \frac{\mathrm{Cov}(Y_t, Y_{t-s})}{\mathrm{Var}(Y_t)} = \frac{\gamma_s}{\gamma_0}, \qquad s = 1, 2, \ldots$$

PSA: Population vs. Sample Quantities

- It is important to note that the ACF is a property of a *process*, not of a *sample* (i.e., the observed time series). This makes them *population quantities* or *parameters*.
- The statement on the previous slide says that the random variables $\{Y_1, Y_2, \ldots\}$ generated by a white noise process are uncorrelated.
- Population quantities are *unobserved*. The best we can hope for is to *estimate* them from a sample (a time series).

PSA: Population vs. Sample Quantities

- To use the normal distribution as an analogy: it has two parameters, μ and σ^2. These are *parameters* and thus *unobserved*.
- In a simulation exercise, I can *pretend* to know what μ and σ^2 are.
- E.g., I can set $\mu = 0$ and $\sigma^2 = 4$, simulate 1000 random numbers y_t, and give them to you.
- Unlike me, you won't know what μ and σ^2 are. At best, you can *estimate* them, based on the *sample mean and variance*

$$\bar{y} \equiv \frac{1}{N}\sum_{i=1}^{N} y_i \quad \text{and} \quad s_y^2 \equiv \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2.$$

- If the sample is large enough, then these will be *close* to $\mu = \mathbb{E}[Y]$ and $\sigma^2 = \mathrm{Var}(Y)$ by the law of large numbers (*LLN*).

The Correlogram

- Applying the analogy to the ACF, the *Sample ACF* or *correlogram* is defined as

$$\hat{\tau}_s = \frac{\hat{\gamma}_s}{\hat{\gamma}_0} = \frac{\sum_{t=s+1}^{T}(y_t - \bar{y})(y_{t-s} - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2}, \qquad \bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t.$$

- The correlogram is a *sample quantity*, i.e., I can compute it from a given time series.
- If I want to test if the time series is white noise, I can compare my SACF to the ACF of a white noise process.
- If the two are significantly different, then I can reject the null that the time series was generated by a white noise process.
- See exercises and the spreadsheet `simulations.xlsx`.

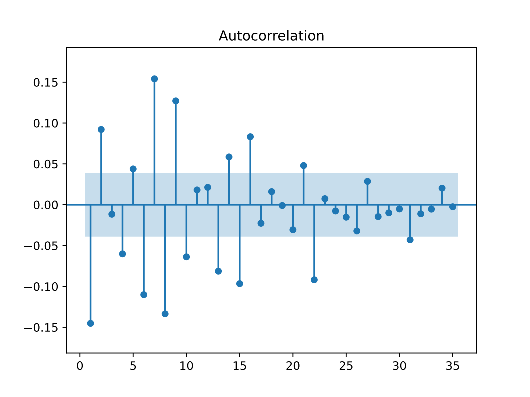
Testing if an Autocorrelation is Zero

- One can show that under the null that the data were generated by a white noise process, the sample autocorrelations are asymptotically[4] normally distributed with zero mean and variance $1/T$.
- This implies that a sample autocorrelation is significantly different from zero if its absolute value is larger than $1.96/\sqrt{T}$.
- We can also test whether the first m autocorrelations are zero jointly: under $H_0 : \tau_s = 0, s \geq 1$, the *Ljung-Box* Q-statistic

$$Q(m) = T(T+2)\sum_{s=1}^{m} \frac{\hat{\tau}_s^2}{T-s} \quad \xrightarrow{d} \quad \chi^2(m).$$

[4]Formally: under $H_0 : \tau_s = 0, s \geq 1$, $\sqrt{T}\hat{\tau}_s \xrightarrow{d} N(0,1)$.

Example: Correlogram of S&P500 returns



- The shaded blue area represent the critical value of $1.96/\sqrt{T} = 0.039$, so, e.g., the autocorrelations at lags 1 (-0.145) and 2 (0.092), are significant, while the one at 3 (0.0116) is not.

Example: Q-stats for S&P500 returns

	lb_stat	lb_pvalue
1	53.146739	3.095389e-13
2	74.532185	6.539456e-17
3	74.874486	3.854848e-16
4	84.009633	2.460691e-17
5	88.860498	1.165864e-17
6	119.473478	2.102105e-23
7	179.461709	2.531020e-35
8	224.529615	4.248635e-44
9	265.378553	5.617980e-52
10	275.677251	2.125260e-53

E.g., $Q(5) = 88.86$ can be used to test the null that the first 5 autocorrelations are jointly zero. The critical value is 11.07, so the test rejects.

- The Q statistics confirm the presence of autocorrelation (p-values less than 5%).
- Conclusion: some autocorrelation, and hence predictability, in the returns; returns are not white noise. Unclear if predictability is sufficient to exploit with a trading strategy.

Outline

From Returns to Asset Prices

- We have seen that the EMH suggests that white noise is a reasonable model for stock returns:
  $$r_t = U_t, \quad \text{where } U_t \text{ is white noise (not necessarily normal).}$$
- Recall the definition of log returns:
  $$r_t = \log P_t - \log P_{t-1}.$$
- Putting the two together implies that
  $$\begin{aligned} \log P_t - \log P_{t-1} = U_t &\Leftrightarrow \\ \log P_t &= \log P_{t-1} + U_t. \end{aligned}$$
- This characterizes the log asset price as a *random walk*.

Definition

A *random walk* is the stochastic process

$$Y_t = Y_{t-1} + U_t,$$

where U_t is white noise and Y_0 is some fixed starting value.

Properties of the Random Walk

- The random walk behaves very differently from white noise.
- A quick calculation shows that

$$Y_t = Y_0 + U_1 + U_2 + \cdots + U_t = Y_0 + \sum_{s=1}^{t} U_s.$$

- From this, it is immediate (see exercises) that

$$\mathbb{E}(Y_t) = Y_0 \quad \text{and}$$
$$\mathrm{var}(Y_t) = \sigma^2 t.$$

- One can also show that

$$\mathrm{corr}(Y_t, Y_{t-k}) = \sqrt{(t-k)/t}.$$

Properties of the Random Walk

- In words:
  - The effect of a "shock" U_t is permanent; U_t is in all future values Y_s, $s \geq t$, whereas for a white noise process, U_t only affects Y_t.
  - The variance increases over time, because we add up more and more of the U_t, all of which are random.
  - The correlogram decreases slowly, approximately linearly (see also `simulation.xlsx`).
- We say that a random walk is not *mean reverting*; one can show that it will (eventually) hit each and every level L, and its excursions can take arbitrarily long.

Outline

Stationary Processes

- Earlier, we rejected the null that the returns on the S&P500 are white noise (although they are close).
- This also implies that the log stock prices are not (exactly) random walks.
- This means that we need to generalize these concepts to allow for other types of stochastic process.
- Specifically, instead of pure white noise, we will consider *stationary processes*.
- Similarly, we generalize the concept of a random walk to *integrated processes*.
- Specific instances of these processes (ARMA and ARIMA models, respectively) will be considered in Weeks 3 and 4.

Stationarity

Definition

A process $\{y_t\}$ is called *weakly stationary* (or second-order, covariance stationary) if the first two moments are time-invariant:

$$\mathbb{E}[Y_t] = \mu, \quad \mathrm{var}(Y_t) = \gamma_0 = \sigma^2, \quad \mathrm{cov}(Y_t, Y_{t-s}) = \gamma_s, \quad t \in \{0, \ldots, T\}, s \geq 1.$$

- This means that the mean, variance, and autocovariances (or autocorrelations) do not change over time; i.e., the autocovariance γ_s depends only on the lag s, not on t.
- Intuitively, there should not be a significant difference if I calculate the mean, variance, and ACF from the first or second half of the sample.
- White noise is one example of a *stationary* process.
- As we saw, the random walk is not stationary; its variance changes over time.

Integrated Processes

- Recall that if returns are white noise, then log prices follow a random walk:

$$\log P_t = \log P_{t-1} + U_t$$

- Alternatively, if log prices follow a random walk, then returns are white noise:

$$r_t = \log P_t - \log P_{t-1} = U_t$$

- We write $\Delta \log P_t$ for $\log P_t - \log P_{t-1}$.
- So a process Y_t is a random walk if ΔY_t is white noise.

Integrated Processes

- We saw above that for the S&P500, we did not get white noise after differencing, but some other stationary process.
- Such processes are called *integrated*.

Definition

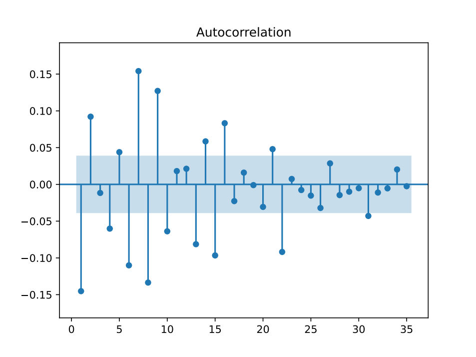
A process Y_t is called *integrated* of order 1, or *I(1)*, if it is non-stationary itself, but $\Delta Y_t = Y_t - Y_{t-1}$ is stationary.

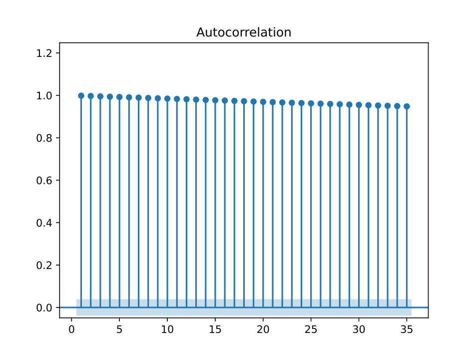
- The random walk is the simplest example of an I(1) process.
- A stationary process is also called *I(0)*.
- An I(2) process would need to be differenced twice to be stationary, but this is rarely necessary in practice.

Properties of Integrated Processes

- Integrated processes have correlograms that stay close to one, which *die out very slowly*.
- An informal way to check whether the stationarity assumption is reasonable is by inspecting the graph and the correlogram of the series. If the graph displays a tendency to revert to a constant mean, with a more or less constant variance, and the correlogram converges to zero *exponentially fast*, then stationarity may be assumed. A formal test will be introduced later.
- Besides prices, many financial and economic time series (e.g., GDP) do not seem to be stationary, because they display a trending mean, and a variance that increases with the level of the process. The latter phenomenon is usually dealt with by a log-transformation, but then quite often the series is still not stationary.

Example: ACF of S&P500 returns and log prices





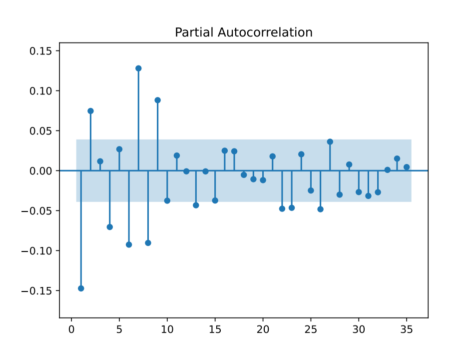
Note: Partial Autocorrelation Function

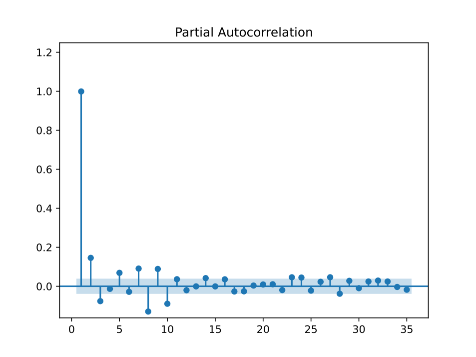
- A concept related to the ACF is the (sample) *partial autocorrelation function* ((S)PACF) $\hat{\tau}_{kk}, k = 1, 2, \ldots,$ where $\hat{\tau}_{kk}$ is the OLS estimator of τ_{kk} in the regression

$$y_t = \alpha + \tau_{k1} y_{t-1} + \ldots + \tau_{kk} y_{t-k} + e_t.$$

  Note: this is not the model for y_t, just a regression to estimate τ_{kk}!
- The PACF measures the correlation between y_t and y_{t-k}, *controlling* for the effect of the intermediate lag. I.e., τ_{kk} only measures the *direct* effect of y_{t-k} on y_t.
- For a random walk, it drops to zero after the first lag, because only y_{t-1} has a direct effect.
- For a stationary process, the ACF and PACF converge to zero at a geometric (exponential) rate as k increases.
- If the sample ACF and PACF of a time series do not seem to converge at all, or too slowly (linearly), then this is an indication of nonstationarity.

Example: PACF of S&P500 returns and log prices





More Properties of Integrated Series

- *No mean-reversion*. Like the random walk, I(1) processes do not revert to a mean.
- *Persistence of shocks*. Also, the effect of past shocks u_{t-i} does not die out, whereas for stationary series the effect will decay exponentially. Important for economic policy.
- *Increasing forecast intervals*. For I(0) time series, the long-run 95% forecast interval converges to the unconditional mean $\pm$ twice the unconditional standard deviation. For an I(1) process the forecast variance does not converge, so forecasts intervals keep increasing.
- *Spurious regressions*. When regressing two integrated time series onto each other, the R^2 and t-statistic may become very large even if they are totally independent. This is avoided if we regress Δy_t on Δx_t.
- *Asymptotic properties of estimators and tests*. In regressions with I(1) variables, the usual statistical theory breaks down (asymptotic normality of estimators, t-tests, etc).

Outline

Learning Goals

Students

- know the definitions of simple and log returns,
- know the definition of white noise,
- understand the ACF and PACF and their sample analogs,
- are able to use the correlogram and Q-statistics to test if a series was generated by a white noise process, and
- are able to distinguish stationary and integrated processes.

Homework

- Exercise 2
- Questions 9b and 12b from Chapter 6 of Brooks (2019)