

**Bản quyền dịch thuật và diễn giải thuộc về justinnguyen92© - telegram**

## Mean Reversion

### 1. Định nghĩa

Hồi quy về trung bình (*mean reversion*) là một thuộc tính của chuỗi thời gian (*time series*), cho thấy chuỗi này sẽ có một giá trị trung bình dài hạn mà xung quanh đó các giá trị có thể dao động theo thời gian, nhưng cuối cùng sẽ quay trở lại mức trung bình này (*Chan, 2013; Ehrman, 2006; Vidyamurthy, 2004*). Thuộc tính này đóng vai trò quan trọng trong giao dịch cặp (*pair trading*), nơi mà chuỗi thời gian hồi quy về trung bình được tạo ra bằng cách kết hợp hai (hoặc nhiều) tài sản.

Hồi quy về trung bình cho phép nhà giao dịch mua ở mức giá thấp với kỳ vọng rằng giá sẽ quay trở lại mức trung bình dài hạn theo thời gian. Khi giá quay lại mức trung bình lịch sử, nhà giao dịch sẽ đóng vị thế để hiện thực hóa lợi nhuận.

Tất nhiên, điều này phụ thuộc vào giả định rằng mối quan hệ lịch sử giữa các tài sản sẽ tiếp tục trong tương lai, điều này tiềm ẩn rủi ro; do đó, việc theo dõi cẩn thận là rất quan trọng.

#### Giải thích:

- Hồi quy về trung bình (*mean reversion*): Đây là chiến lược trong đầu tư và giao dịch cặp tài sản. Nó giả định rằng các tài sản có mối quan hệ chặt chẽ về giá trị sẽ có xu hướng quay trở lại khoảng cách trung bình của chúng sau khi xảy ra biến động.
- Giao dịch cặp (*pairs trading*): Đây là phương pháp kết hợp hai hoặc nhiều tài sản tương quan để tìm kiếm cơ hội thu lợi nhuận từ sự biến động tương đối giữa các tài sản này.
- Rủi ro: Nếu mối quan hệ lịch sử bị phá vỡ hoặc không còn đúng trong tương lai, chiến lược này có thể dẫn đến thua lỗ.

### 2. Các loại Mean Reversion

#### Longitudinal or Time Series Mean Reversion (Hồi quy về trung bình theo chuỗi thời gian):

Xảy ra khi sự hồi quy diễn ra theo chiều thời gian.

Ví dụ: Giá tài sản dao động quanh mức trung bình và cuối cùng trở lại mức trung bình dài hạn sau khi lệch về một phía.

#### Cross-Sectional Mean Reversion (Hồi quy về trung bình theo chiều không gian):

Xảy ra khi sự hồi quy diễn ra trên nhiều tài sản trong cùng một thời điểm. Một số tài sản có thể giảm giá trong khi những tài sản khác tăng, nhưng trung bình chung vẫn ổn định.

### 3. Stationarity (Tính dừng)

#### Định nghĩa:

"Tính dừng (*stationary*) là thuộc tính của chuỗi thời gian mà trong đó trung bình (*mean*), phương sai (*variance*), và hiệp phương sai (*covariance*) không thay đổi theo thời gian. Một chuỗi dữ liệu stationary

có thể được coi là mean-reverting vì các tham số của nó luôn cố định."

#### 4. Unit-Root Stationarity (Tính dừng không có gốc đơn)

Unit Root: Nếu một chuỗi có "gốc đơn" ("unit root"), nó sẽ không có tính dừng (*stationary*) mà có xu hướng phân kỳ (*diverge*) thay vì hồi quy về trung bình. Chuỗi ngẫu nhiên *random walk* là một ví dụ điển hình của chuỗi có gốc đơn *unit root*.

#### 5. Mô hình Mô hình AR(1) (Autoregressive Order 1):

Phương trình mô hình AR(1):

$$y_t = \mu + \rho y_{t-1} + \epsilon_t$$

**Trong đó:**

- $y_t$ : Giá trị tại thời điểm  $t$ .
- $\mu$ : Hằng số trôi dạt (drift).
- $\rho$ : Hệ số hồi quy (nếu  $|\rho| < 1$ , chuỗi hồi quy về trung bình).
- $\epsilon_t$ : Thành phần sai số ngẫu nhiên.
- Nếu  $|\rho| \geq 1$ , chuỗi không hồi quy về trung bình và có thể phân kỳ vô tận.

Mặc dù Hồi quy về trung bình (*mean reversion*) và tính dừng không có gốc đơn (*unit-root stationarity*) không phải là các khái niệm tương đương nhau, nhưng trong thực tế, tính dừng không có gốc đơn (*unit-root stationarity*) là một đại diện thuận tiện cho Hồi quy về trung bình (*mean reversion*) (Tsay, 2010, 2013). Trên thực tế, việc kiểm tra tính dừng không có gốc đơn là phương pháp tiêu chuẩn để xác định tính hồi quy về trung bình trong thực tế.

Lấy sai phân (*differencing*) là một thao tác thường được sử dụng để đạt được tính dừng (Tsay, 2010, 2013). Nó đề cập đến việc lấy sự chênh lệch giữa các mẫu liên tiếp của chuỗi thời gian  $y_1, y_2, y_3, \dots$  để tạo ra:

$$\Delta y_t = y_t - y_{t-1}$$

Ý nghĩa của thao tác này là một chuỗi thời gian không có tính dừng, chẳng hạn như một chuỗi bước ngẫu nhiên (*random walk*), có thể trở thành chuỗi dừng sau khi lấy sai phân (*differencing*). Đây chính là trường hợp khi lấy sai phân *log-price* của một tài sản để tính toán *log-return*. Khi đó, chúng ta nói rằng chuỗi *log-price* được tích hợp ở bậc 1 (*integrated of order 1*). Ngoài ra, cũng có thể xem xét sai phân ở bậc cao hơn.

**Ví dụ code:**

Lấy sai phân *log-price* của một tài sản để tính toán *log-return* (ví dụ: chuỗi *time-series* của EURUSD với giá đóng cửa *close*) - kiểm tra tính dừng của ví dụ EURUSD trên.

```
#
=====
=====
# VÍ DỤ STATIONARY CỦA MỘT CHUỖI TIME SERIES
# ví dụ code Lấy sai phân log-price của một tài sản để tính toán log-return (log
lợi nhuận)
# (ví dụ: chuỗi time series của EURUSD với giá đóng cửa close) - kiểm tra tính
```

```

dùng của ví dụ EURUSD
#
=====
=====
# Import các thư viện cần thiết
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.stattools import adfuller

# Tạo chuỗi thời gian giả lập giá đóng cửa của cặp EUR/USD
np.random.seed(42)
T = 200 # số lượng điểm dữ liệu
initial_price = 1.1 # giá ban đầu
returns = np.random.normal(0, 0.001, T) # lợi nhuận ngẫu nhiên nhỏ
prices = initial_price * np.exp(np.cumsum(returns)) # chuỗi giá giả lập

# Tạo DataFrame chứa giá đóng cửa EUR/USD
df = pd.DataFrame({'Close': prices}, index=pd.date_range(start='2022-01-01',
periods=T))

# Tính toán log-price và log-return (lấy sai phân)
df['Log_Close'] = np.log(df['Close']) # log-price
df['Log_Return'] = df['Log_Close'].diff() # log-return (sai phân)

# Kiểm tra tính dừng bằng Augmented Dickey-Fuller (ADF) Test cho log-return
adf_result = adfuller(df['Log_Return'].dropna())
p_value = adf_result[1]

# Vẽ biểu đồ log-price và log-return
plt.figure(figsize=(14, 6))

plt.subplot(2, 1, 1)
plt.plot(df['Log_Close'], label='Log-Price (EUR/USD)', color='blue')
plt.title('Log-Price of EUR/USD')
plt.xlabel('Time')
plt.ylabel('Log-Price')
plt.legend()

plt.subplot(2, 1, 2)
plt.plot(df['Log_Return'], label='Log-Return (Differencing)', color='orange')
plt.title('Log-Return (First Differencing)')
plt.xlabel('Time')
plt.ylabel('Log-Return')
plt.legend()

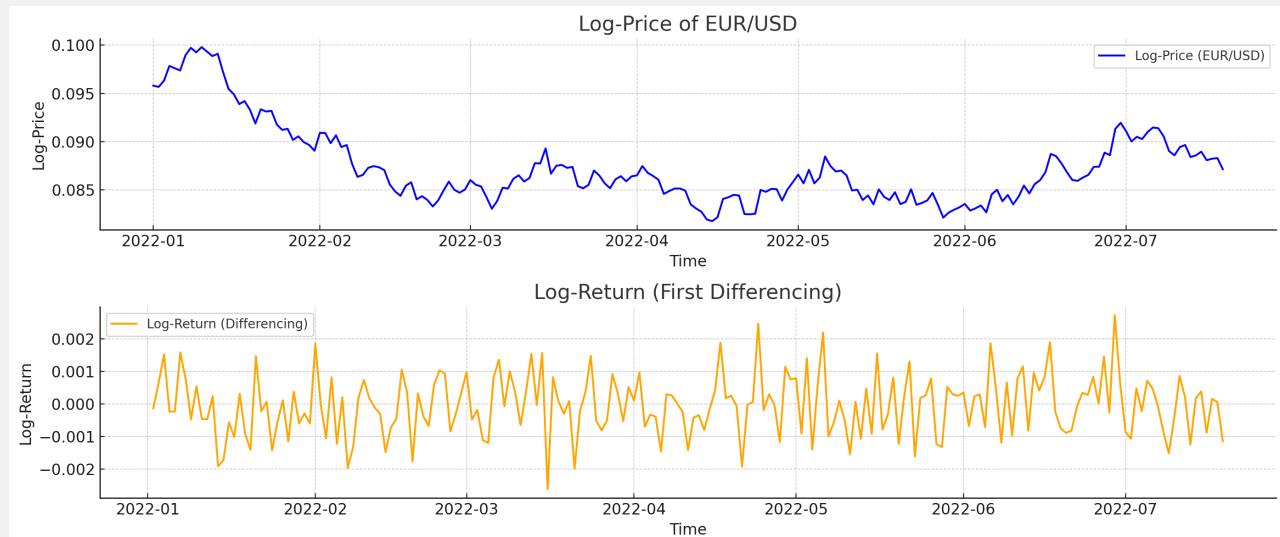
plt.tight_layout()
plt.show()

p_value

```

## Result

3.0415339027305817e-27



## Phân tích kết quả kiểm tra tính dừng cho log-return của EUR/USD

### 1. Biểu đồ Log-Price và Log-Return:

- **Log-Price (trên):** Chuỗi **log-price** biến động dần theo thời gian, có xu hướng tăng hoặc giảm, không cố định quanh mức trung bình → Có thể là chuỗi không dừng.
- **Log-Return (dưới):** Chuỗi **log-return** là sai phân bậc 1 của **log-price**. Biểu đồ cho thấy chuỗi này dao động quanh mức 0 mà không có xu hướng tăng hoặc giảm dài hạn rõ ràng.

### 2. Kết quả kiểm tra ADF Test (Augmented Dickey-Fuller):

- **P-value:**  $\approx 3.04 \times 10^{-27}$ , nhỏ hơn 0.05  $\rightarrow$  Bác bỏ giả thuyết  $H_0$ .
- **Kết luận:** Chuỗi **log-return** có tính dừng  $\rightarrow$  Đây là một chuỗi *stationary* sau khi lấy sai phân bậc 1.

### Tóm lại:

- Chuỗi giá **log-price** ban đầu không có tính dừng.
- Sau khi lấy sai phân **log-price** (**log-return**), chuỗi trở thành dừng, phù hợp với lý thuyết *mean reversion* trong chuỗi *time-series*.