暗号通貨に対する統計的手法による投資戦略の検討

A Study on Investment Strategy for Cryptocurrency with Statistical Methods

統計解析研究室
5CS23 小枩谷勇二

目次

1. 研究背景
2. 研究目的
3. 研究方法
4. 開発環境
5. 時系列解析について
6. BTC価格データの統計的性質の調査結果
7. 解析手法・投資戦略の検討
8. 結果
9. 考察
10. 今後の展望

研究背景

研究背景

近年･･･

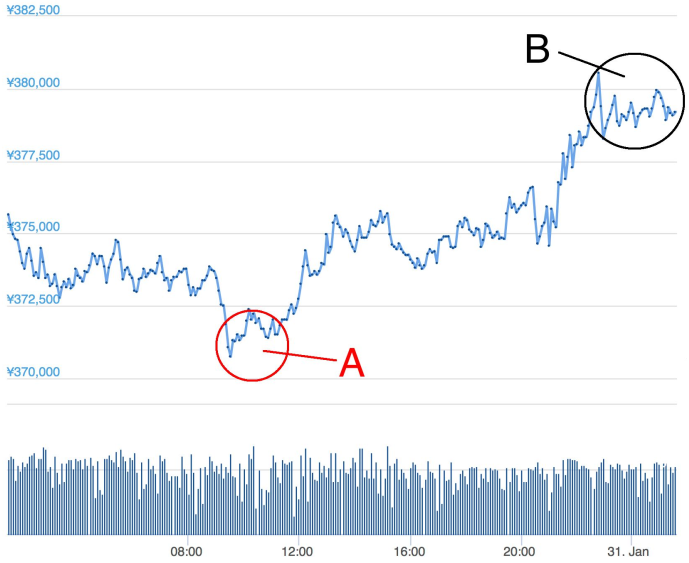
暗号通貨

･BTC(ビットコイン)
･XRP(リップル)
･その他

これらをはじめとした**暗号通貨**が開発されている.

研究背景 - 暗号通貨の価格

暗号通貨の価格：従来の通貨同様，**常に変動**している．



BTC/JPY 2019/01/30-2019/01/31

価格帯Aで購入し，Bで売却する

↓

利益を得ることが可能．

↓

<u>暗号通貨に対し，投資戦略を検討したい．</u>

研究背景 - 暗号通貨のデータ

暗号通貨 : 日時に対する**価格の数値データが存在**する時系列データ

⇩

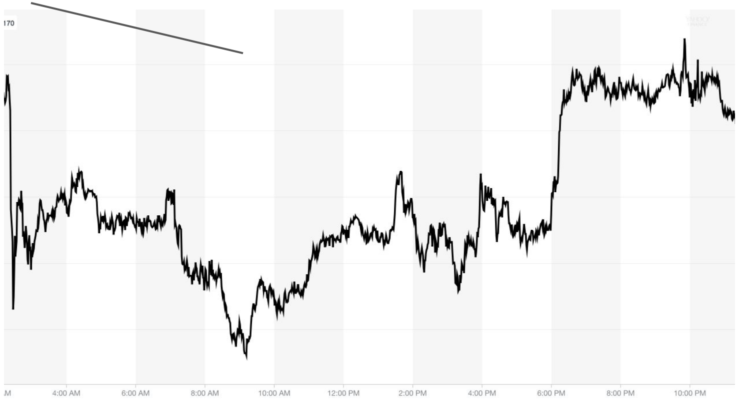
統計的手法による，価格データの解析・投資戦略の検討

研究背景 - 暗号通貨のデータ

暗号通貨 : 日時に対する**価格の数値データが存在**する時系列データ

同等のものには, **株式･外国為替**等が存在する.

P&G 2014/02-2019/01



ドル円 2019/02/06-2019/02/07

研究背景 - 先行研究について

株式:

データマイニングを利用してその価格を予測するシステムの開発[1]などが行われている.

外国為替:

多変量解析による価格の変動を要因を分析する研究[2]などが行われている.

暗号通貨:

統計的手法による解析等を行う試みは**<u>ごく少数</u>**にとどまっている.

研究目的

研究目的

暗号通貨の価格データに対して，**<u>統計的手法による投資戦略を検討すること</u>**が，本研究の目的である.

研究方法

研究方法

本研究は大きく分けて以下の手順で行う.

1. API(Cryptowatch[3])によるBTC価格データの取得
2. 価格データに対する統計的性質の調査
3. 価格データ解析手法・投資戦略の検討
4. 投資戦略による売買シミュレーション

研究方法 - 価格データについて

本研究において，使用するBTC価格データは以下の通り.

・取引所：bitFlyer

・期間：2018/10/31/ 12:00:00 - 2018/11/30 23:55:00

・間隔：開始から5分ごと

研究方法 - 統計的性質の調査について

暗号通貨 : 日時に対する価格の数値データが存在する**時系列データ**

↓

時系列解析を行う.

開発環境

開発環境

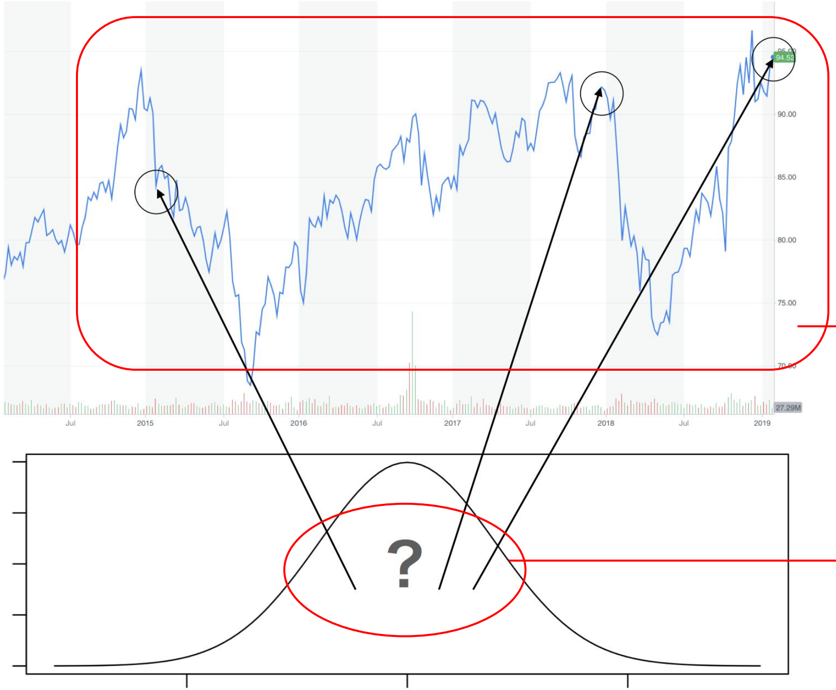
開発環境は以下の通り.

用途	環境	外部依存関係
開発OS	macOS Sierra10.12.6	-
価格データ取得	Python3.6.4	mysqlclient
価格データ保存	MySQL5.7.21	-
統計的性質調査	R3.4.4	tseries urca RMySQL
価格解析	Python anaconda3-5.1.0	TA-Lib

時系列解析について

時系列解析について - 考え方

時系列データは，**全ての時点のデータは1度のみ観測可能.**



観測**データは確率変数によって算出された**ものとみなす.

確率分布の性質・構造を推定

時系列解析について - 使用するデータ

本研究においては対数差分を解析する.

t 時点での価格をP_t, $\ln(x)$をxの自然対数としたとき, 対数差分は,

$$\ln(P_t) - \ln(P_{t-1})$$

となる.

時系列解析について - 使用するデータ

対数差分式は

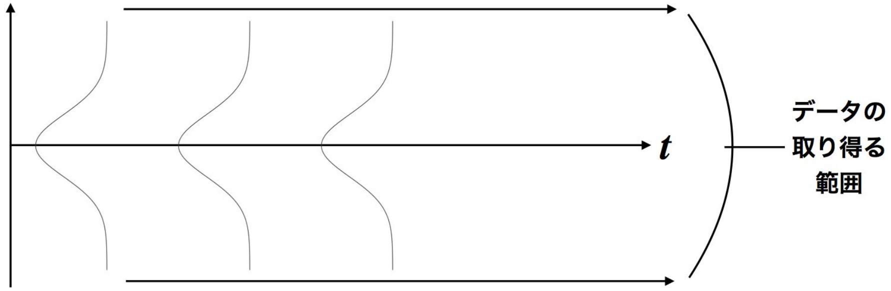
$$\frac{P_t - P_{t-1}}{P_{t-1}}$$

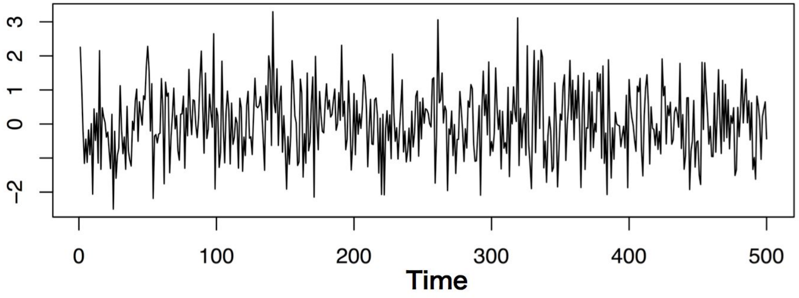
と変形可能.

すなわち, 原系列データの**対数差分は現時点のデータと直前のデータの変化率**とみなせる.

時系列解析について - 定常性について

時系列解析：解析するデータが**定常性を有している必要がある**.





左の図では，**確率分布の平均が一定である**ことが確認できる.

予測される**データの範囲は時点に関係なく一定である**と言える.

時系列解析について - 定常性について

解析対象のデータが定常性を有しているか確認する必要がある.

単位根検定を行う.

本研究では，検定結果の**p値が0.1未満**ならば，**データは定常性を有する**ものとみなす.

時系列解析について - 自己相関

過去データを元に将来の価格変動を予測したい.

⇩

現在のデータと過去データの関係を調査

⇩

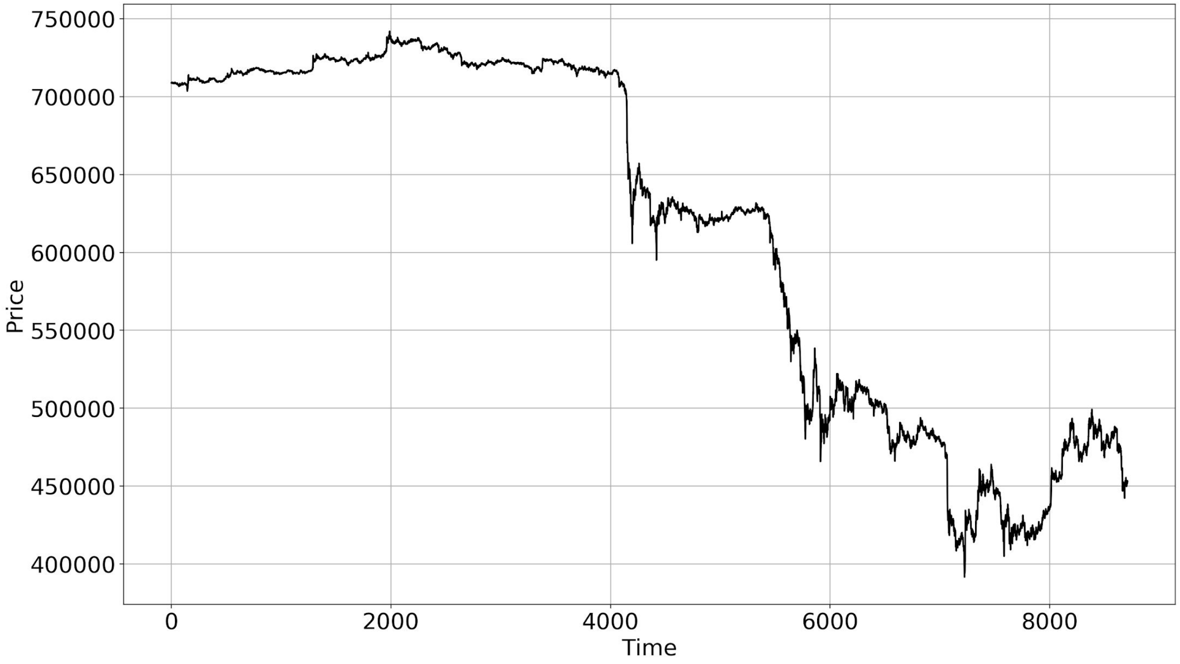
自己相関を調査

※自己相関:現時点のデータと，過去の時点のデータとの関係を示した指標.

BTC価格データの統計的性質

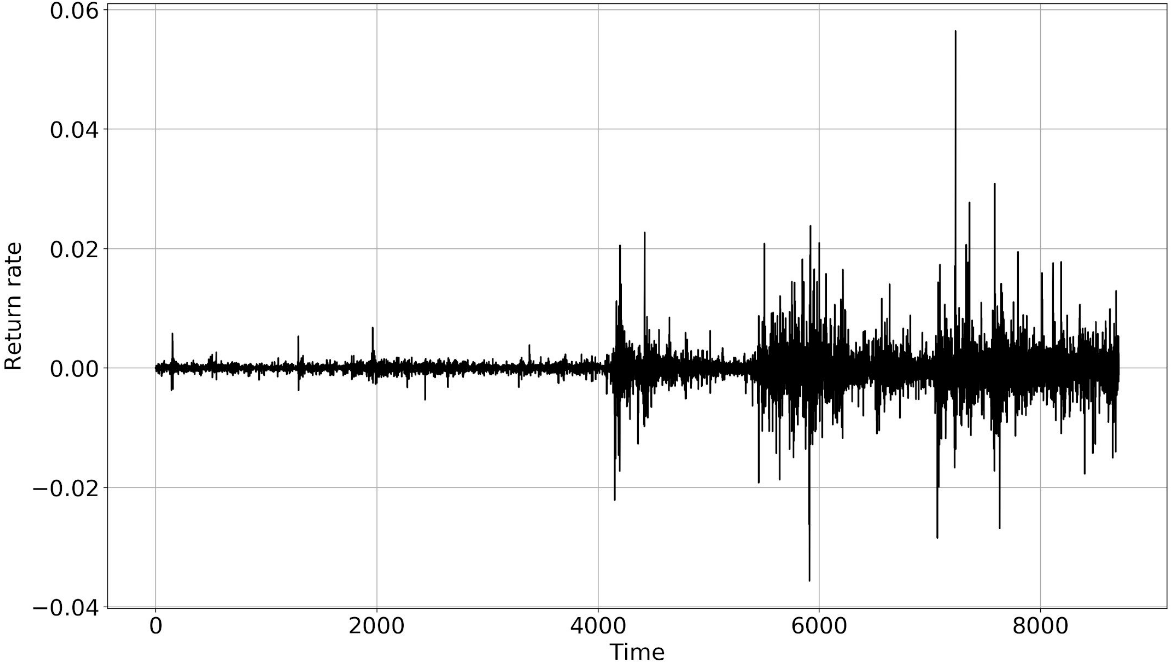
BTC価格データの統計的性質 - 対数差分

2018/10/31 12:00:00 - 2018/11/30 23:55:00のBTC価格データを示す.



BTC価格データの統計的性質 - 対数差分

同期間のBTC価格データの対数差分は以下の通り.



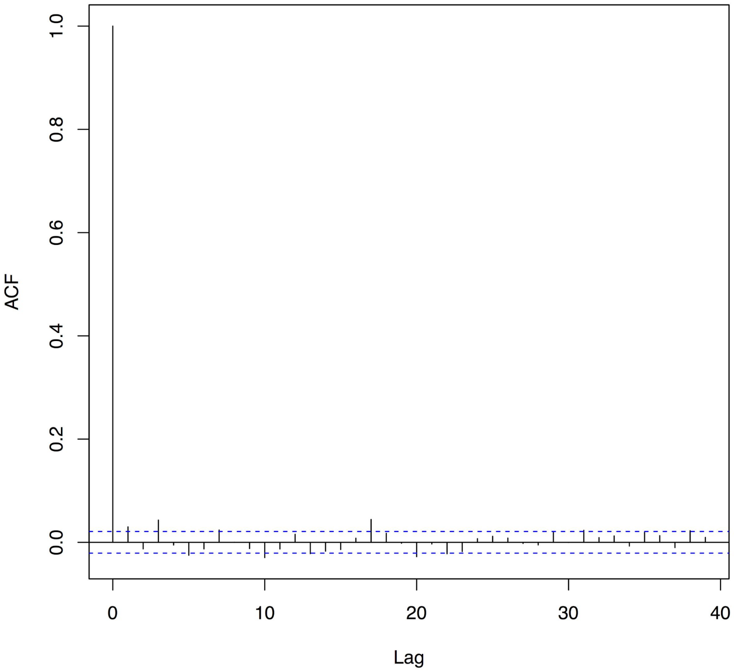
BTC価格データの統計的性質 - 定常性

単位根検定を行った結果は以下の通り.

対象データ	p値	結果
原系列データ	0.5629	定常性を有さない
対数差分	0.01未満	定常性を有する

BTC価格データの統計的性質 - 自己相関

BTCの対数差分に対する，自己相関係数の算出結果は以下の通り.



自己相関係数は，いずれの時点においても**0.05未満**である.

有意な自己相関とは言えない.

解析手法・投資戦略の検討

解析手法 - 方針

BTC価格の対数差分は**定常過程ではあるが自己相関はない.**

↓

ほぼランダムに発生している.

↓

確率分布の性質・構造の推定は困難

↓

対数差分の本解析は**有効ではない.**

解析手法 - 方針

価格データそのものを解析する.

価格データ: 定常過程ではない.

⇒ 解析手法は限られてくるが, **不可能ではない.**

本研究では, **<u>単純移動平均線</u>**を利用して解析する.

解析手法 - 単純移動平均

ある時系列データ$X_1, X_2, \ldots$ に対する期間p の，任意の時点t での，

単純移動平均$SMA(t, p)$の値は，

$$SMA(t, p) = \frac{1}{p}\sum_{i=0}^{p-1} X_{t-i}$$

と算出される.

単純移動平均の算出例(期間=5時点)は以下の通り.

時点	1	2	3	4	5	6
実現値	10	20	30	40	50	60
移動平均	-	-	-	-	30	40

解析手法 - 単純移動平均

時系列データに対する単純移動平均線(期間=13時点)の描画例を示す.

※赤線:単純移動平均線, 青線:原系列データ

単純移動平均線による投資戦略

単純移動平均線を利用し，**買い時・売り時の目安を示す**ことが可能．

買い時：ゴールデンクロス(GC)
売り時：デッドクロス(DC)
といったものが存在する．

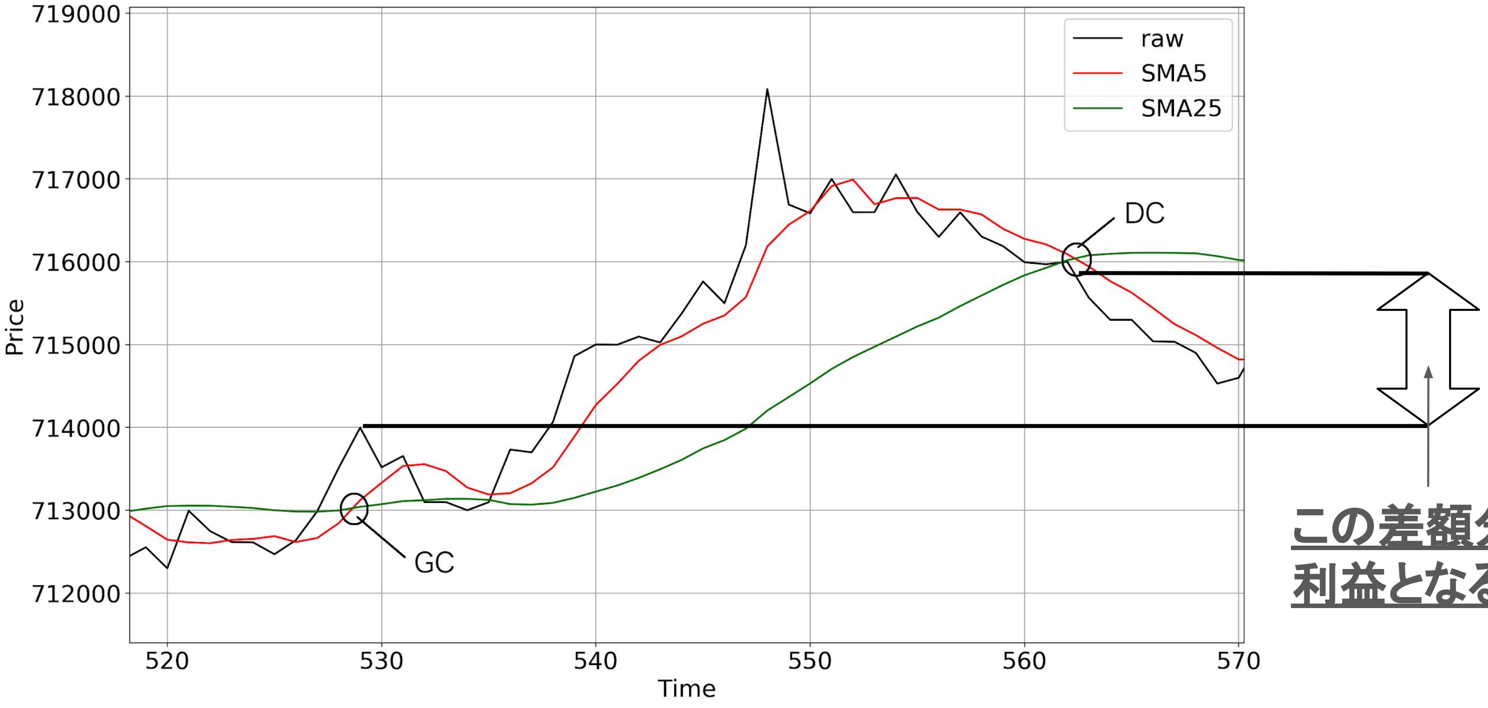
2つの**期間の異なる**単純移動平均線を描画し，

・短期線が長期線を下から上に突き抜けた：GC
・長期線が短期線を上から下に突き抜けた：DC

である．

単純移動平均線による投資戦略

図中において，**GC時点で購入**し**DC時点で売却**した場合，



<u>この差額分が利益となる．</u>

売買・シミュレーションのルール

・2つの単純移動平均線の期間：5時点・25時点

・GC時点で購入，DC時点で売却したと仮定する．

・取引量：1BTC固定

・手数料等は計算に含まない．

・ロスカット等は行わない．

※ロスカット：証券・通貨等を保有している状態で，損失が一定以上の水準に達した場合，その注文に対する決済を行い損失の拡大を防ぐこと．

結果

結果

最終的な売買結果は以下の表の通り.

期間	2018/10/31 12:00:00 - 2018/11/30 23:55:00
売買回数	231
利確回数	77
損益合計	-19424円

考察

考察

最終結果として，利益をあげることはできていない．

本研究における期間：価格は右肩下がり(下降トレンド)

⇨ 購入→売却では**損失が発生しやすい**．

考察

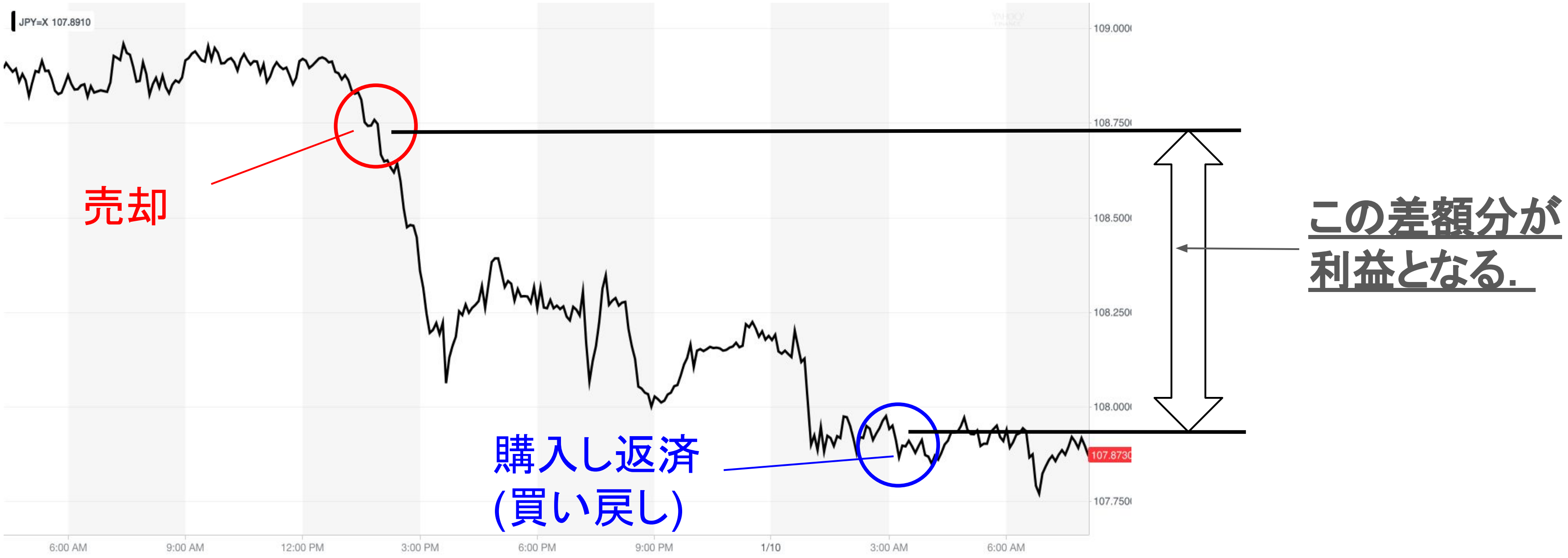
投資家のリスク許容度により，取りうる戦略も異なる．

⇨ 本研究：リスクを**低く抑える**方針で考察

考えられる戦略：**空売りによるリスクヘッジ**

考察 - 空売りとリスクヘッジ

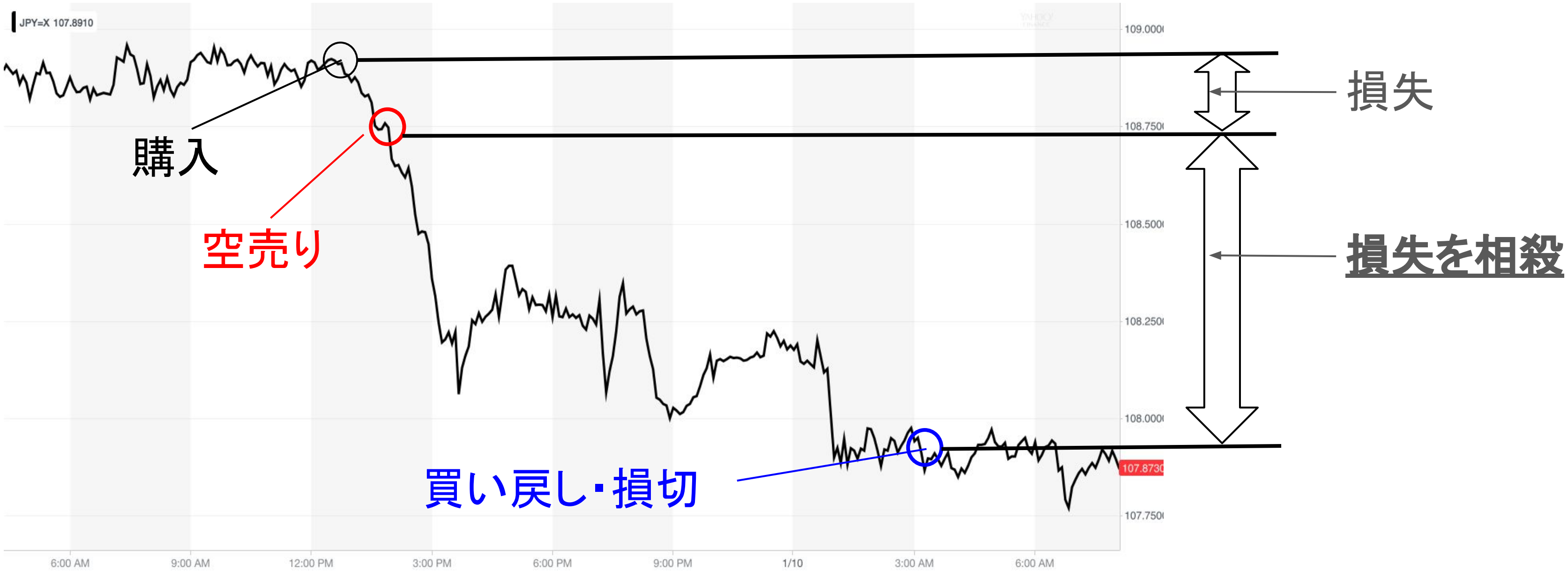
空売り: 事前に取引会社から**通貨等を借り**, 売却すること.



ドル円 2019/01/09-2019/01/10

考察 - 空売りとリスクヘッジ

価格下降時に空売りをした場合



ドル円 2019/01/09-2019/01/10

考察

空売りの持ち高：**下降トレンド終了時**に返済したい．

⇩

トレンドを<u>どのように判断する</u>か．

⇩

<u>**移動平均線の傾きを観測**</u>する．

考察 - まとめ

・古典的な時系列解析(対数差分解析)：**有意ではない.**

・単純移動平均線：下降トレンド中は損失が出やすい.

⇒ **空売り**によるリスクヘッジの必要性.

今後の展望

今後の展望

以下の事柄が，今後の展望の一例として挙げられる．

・他のテクニカル指標の導入・組み合わせによる投資戦略の検討

・移動平均線係数・期間調整

参考文献

[1] 矢部大輔・木村昌臣(2008)「データマイニングによる株価予測システムの開発」

[2] 砂田吉一・橋本次郎(1989)「多変量解析による為替レートの要因分析 (第2回計算機統計学会シンポジウム報告 経済と経営の実証分析 : 計算機の利用)」

[3] Market Data REST API https://cryptowatch.jp/docs/api

参考 - 単位根･定常性 - 単位根とは

単位根：離散確率過程(時系列データ)の<u>**性質を表す**</u>言葉

定義：離散確率過程X_t $(t=1,2,...)$ において，X_tは

$$X_t = \phi_p X_{t-p} + \cdots + \varepsilon_t \quad (p = 1, 2, ..., t-1)$$

であるとみなせる場合，

その係数Φ_p $(p=1,2,...,t\text{-}1)$のいずれかが**1**ならば，
X_tは<u>**単位根過程**</u>である(<u>**単位根を有する**</u>)とされる.

ε_t：ホワイトノイズ

参考 - 単位根･定常性 - 単位根過程

例えば，時系列データX_t $(t=1,2,...)$が，

$$X_t = X_{t-1} + \varepsilon_t$$

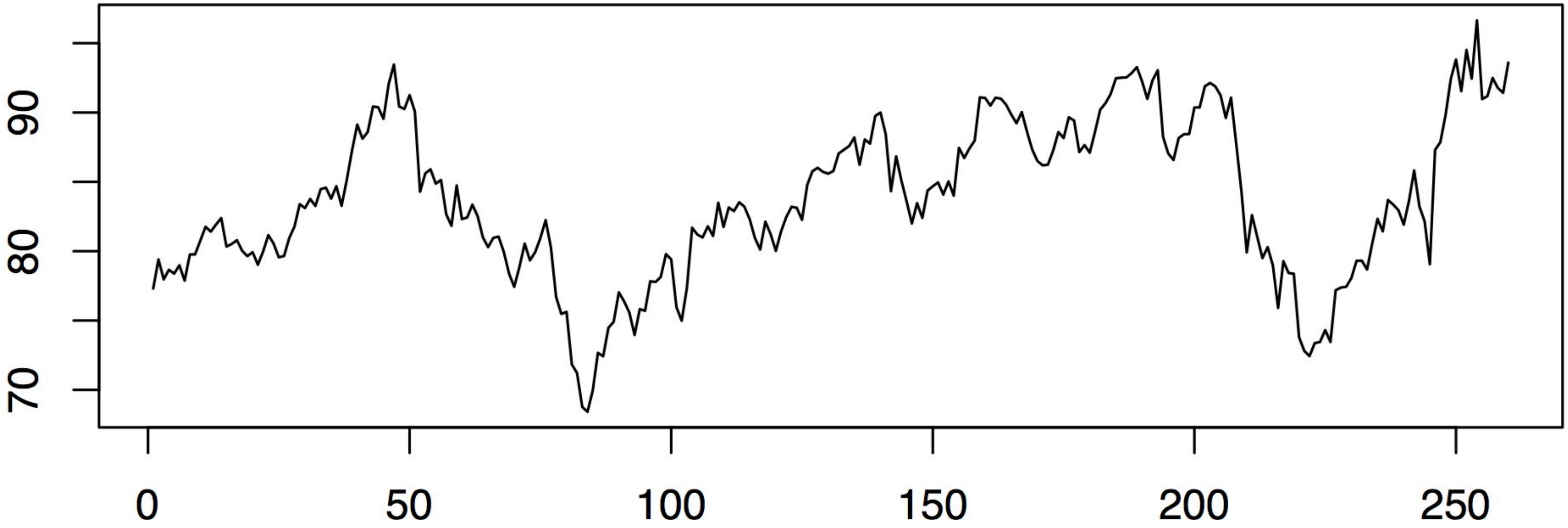
とみなせる場合は，**この時系列データは単位根過程**とされる.

式の解釈：

現在の値は，**直前の値にランダムな変動を加えたもの**である.

参考 - 単位根･定常性 - 単位根過程

単位根過程の一例を以下に示す.



P&G 2014/01-2019/01

参考 - 単位根・定常性 - 単位根過程

単位根過程であるかの判断：**単位根検定**

⇒ **拡張Dickey-Fuller検定**

拡張Dickey-Fuller検定：

帰無仮説：データは単位根過程である.
対立仮説：データは単位根過程ではない.

参考 - 単位根・定常性 - 単位根過程

先ほどの例(P&G 2014/01-2019/01)に対する拡張Dickey-Fuller検定:

⇨ **p値:0.2494**であり**帰無仮説が採択**される.
P&G 2014/01-2019/01の株価は**単位根過程と言える.**
※有意水準10%

※有意水準について:

経済・金融系のデータに対する統計的検定:有意水準10%とされることが多い. 本研究においても同様.

参考 - 単位根・定常性 - 定常過程

時系列解析 : 解析するデータが定常性を有している必要がある.

定常性 : **2種類**存在する.

・**弱**定常性(**弱**定常過程)

・**強**定常性(**強**定常過程)

経済・金融系の時系列データ:

弱定常性を有している(**弱**定常過程である)必要がある.

⇒ 本研究においても同様.

参考 - 単位根･定常性 - 弱定常過程

弱定常過程である時系列データ:どの時点においても,

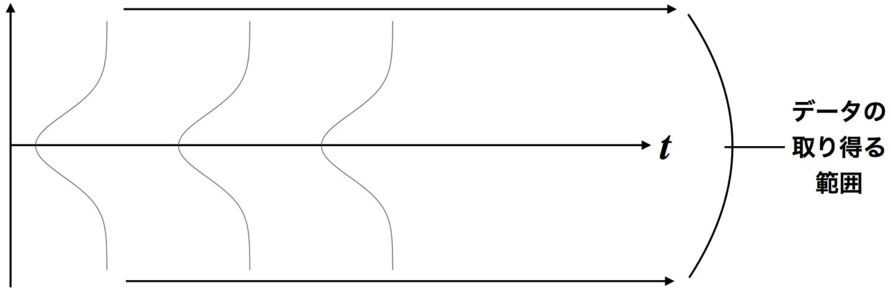
･データの平均が一定
･データの自己共分散が一定

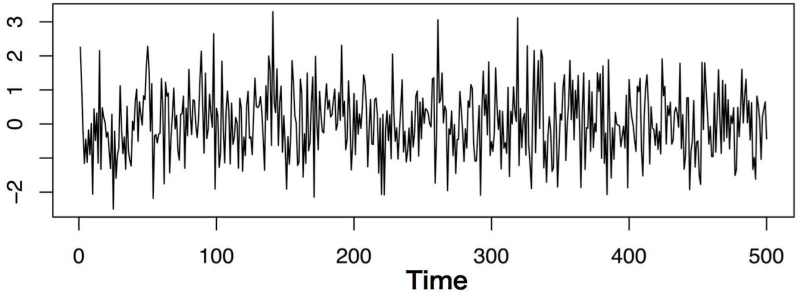
とみなせるデータを指す.

※自己共分散:現時点のデータと過去の任意の時点のデータの関係を示した指標

参考 - 単位根・定常性 - 弱定常過程

すなわち, **弱**定常過程である時系列データとは, 以下の通り.



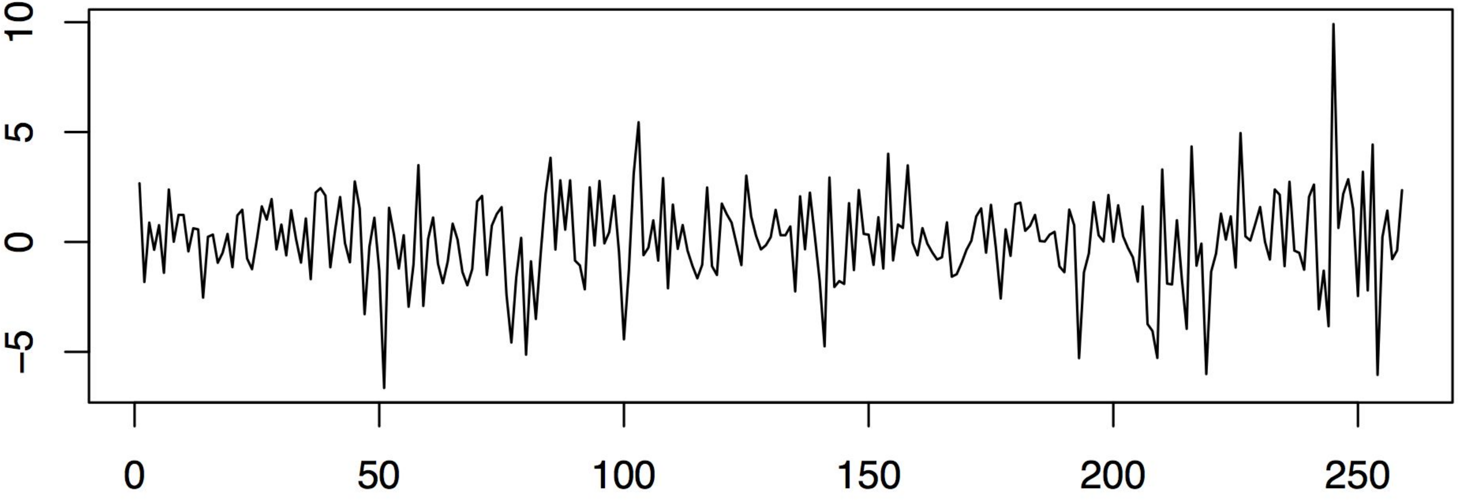


左の図では, **確率分布の平均が一定である**ことが確認できる.

予測される**データの範囲は時点に関係なく一定である**と言える.

参考 - 単位根･定常性 - 弱定常過程

弱定常過程の例：**単位根過程の差分･対数差分**



P&G 2014/01-2019/01の対数差分

※単位根過程は**定常過程ではない**.

参考 - 単位根・定常性 - 定常性の判断

定常過程であるかの判断：**単位根検定**(拡張Dickey-Fuller検定)

⇒ 単位根過程**ではない**必要がある. (**帰無仮説を棄却**)

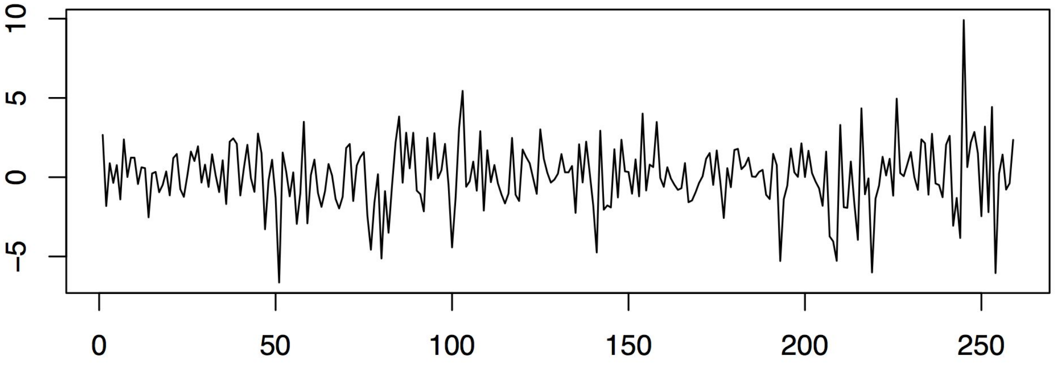
ただし，単位根過程ではない＝定常過程である**とは限らない**.

※拡張Dickey-Fuller検定自体，単位根過程であるとみなせるか否か以上の事象については言及していない.

参考 - 単位根・定常性 - 定常性の判断

データ解析の現場：チャート(時系列)の形で判断

⇨ 確率分布の**平均が一定とみなせる**ならば，**定常過程とする.**



このような時系列データは，(弱)定常過程であるとみなされる.

※そもそも単位根過程の差分・対数差分は弱定常過程.