

回帰分析で株価を予測する

戸山高校 1年F組 川越颯剣
1年F組 福田英輝

動機

コロナショックによって職を失ったり、収入を得られなくなった人が多くいる。そこでどうしてそのような事態になってしまったのか私たちは疑問に感じた。そこで株価の動きを予想することにした。

調査方法 仮の株価予想プログラムをつくり、精度を確認する。

仮説 日経平均株価によってある程度人々の投資の動向が定まる。

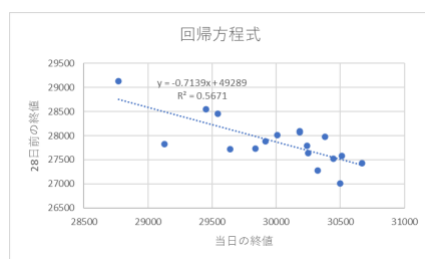
実験①

株予想のグラフをつくる。

目的変数 y を日経平均株価の終値(この日を Y)、説明変数 x を Y の4週間前(この日を X)の日経平均株価の終値として回帰分析をし、回帰方程式のグラフをつくる。ここで、 Y は9月3日から10月1日までとする。(ただし、9月6、20、23日は除く)したがって、 X は8月6日から9月3日となる。(ただし、8月9、23、26日は除く)

結果

日にち	当日の終値	28日前の終値
9月3日	29128	27820
9月7日	29916	27888
9月8日	30181	28070
9月9日	30008	28015
9月10日	30381	27977
9月13日	30447	27523
9月14日	30670	27424
9月15日	30511	27585
9月16日	30323	27281
9月17日	30500	27013
9月21日	29839	27732
9月22日	29639	27724
9月24日	30248	27641
9月27日	30240	27789
9月28日	30183	28089
9月29日	29544	28451
9月30日	29452	28543
10月1日	28771	29128

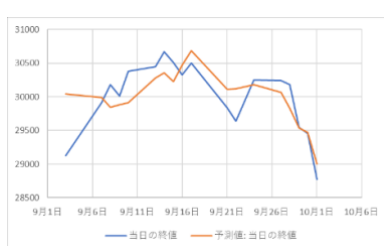


$R^2 = 0.567124409$ から信頼度はあることにはあると言える。

実験②と結果

 実験①の結果が実測値とどれだけ近いかを調べる

日にち	当日の終値	予測値: 当日の終値	差(実測値-予測値)
9月3日	29128	30040.11984	-912.119871
9月7日	29916	29986.10196	-70.10195862
9月8日	30181	29841.5247	339.4753045
9月9日	30008	29885.21563	122.7847333
9月10日	30381	29915.40209	465.5979118
9月13日	30447	30276.05087	170.9491346
9月14日	30670	30354.69454	315.3054585
9月15日	30511	30226.79927	284.2007297
9月16日	30323	30468.29096	-145.2909626
9月17日	30500	30681.18495	-181.1849544
9月21日	29839	30110.02533	-271.025327
9月22日	29639	30116.38037	-477.3803715
9月24日	30248	30182.31396	65.68604142
9月27日	30240	30064.74563	175.2543653
9月28日	30183	29826.43146	356.5685353
9月29日	29544	29538.8657	5.13430041
9月30日	29452	29465.78269	-13.78268746
10月1日	28771	29001.07006	-230.070056



信頼度は実験①の決定係数から $R^2 = 0.567124409$ であるが、予測値と実測値の差の残差平方和の正の平方根(この場合は予測値と実測値の差の絶対値の平均)は255.661795。この値が0になるとき予測値と実測値が一致するから正確性は低い。株価において255円の差は大きいのでもっと正確なプログラムが必要だ。また、実験①と同様に9月のデータを目的変数、10月のデータを目的変数として回帰分析をすると $R^2 = 0.119852322$ で0に近いのでこの仮説は正しくなかったと言える。

新しい仮説

先ほどの仮説が正しくなかったのは、期間が短く短期間の変動が大きい株価に回帰分析が対応できなかったからだと考える。ここで株価チャートにある程度の法則性や類似性が見て取れる長期間においては現時点での株価から未来の株価を回帰分析で予測できると考えた。

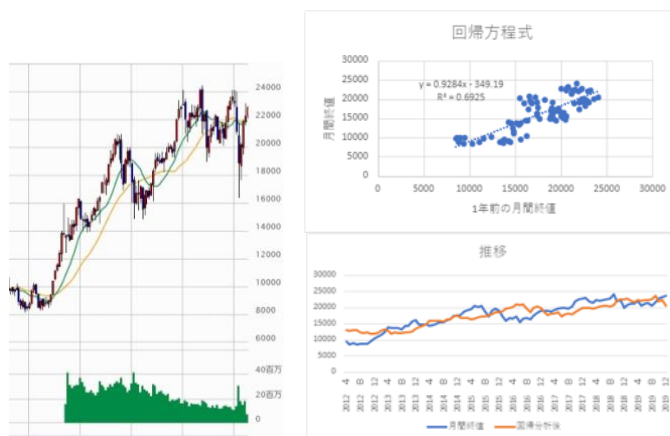
実験③

目的変数 y を日経平均株価の終値(この日を Y)、説明変数 x を Y の1年前(この日を X)の日経平均株価の終値として、回帰分析をし、回帰方程式のグラフを作る。

結果

グラフが安定している時(株価が安定して増えている時)

日経平均株価データの範囲を東日本震災の終わりからコロナショックの前までとした。(2011年から2018年の毎月末の終値が説明変数、説明変数のそれぞれ1年後が目的変数)



↑日経平均株価の推移

↑予想した株価(橙色)と実際の株価(青色)の株価チャート

考察

データの量をたくさんとって決定係数が0.692471361656532であったので正確性は完全とは言えないが担保されているだろう。実際の株価と予想株価の比較のチャートも大まかだが、似たような形をしている。

展望

コロナショックやリーマンショックのときのような株価の動きが変則的な場合に対応した株価チャートをつくる。また、期間が短いときの株価チャートの正確性を向上させる。

反省 主観的意見が多くなってしまっている。データの検定をしなかった。

参考文献 「図解でわかる回帰分析」 涌井良幸 涌井貞美

<<https://m.finance.yahoo.co.jp>>