

一年後の株価はどの程度の精度で予測できるか？

S&P500 の成長率の解析と簡単な確率的未来予測

2020 年 4 月 17 日

はじめに

S&P500（米国株価指数）1950 年 1 月 3 日の終値 16.66 米ドルからさまざまな推移を辿って 70 年 4 ヶ月後の 2020 年 4 月 17 日現在、2,799.55 米ドルをつけている。取引日数にしておよそ 1 万 7 千日。翌日の指数は上がるか、下がるかの二通りしかなく、あとはその上げ幅と下げ幅（まとめて変化率と呼ぶ）がどの程度かという二つの要素、合わせえ「四つの要素の組み合わせ」が 1.7 万回繰り返されて今日の指数に至っている。まず変化率について解析を行ったうえで、「四つの組み合わせ」に関するパターン分析を行い、そこから一年後の指数（以下、単純に株価と呼ぶ）を確率的に予測する簡単な方法を示す。

1) 上昇・下降日数と変化率の解析

過去 70 年間、前日に対して S&P500 の上昇率と下降率の平均値、上昇した日数の合計と下降した日数の合計を集計したものが表 1 である。17,680 日あった取引日の中で前日より指数が高かった日は計 9,388 日で、低かった日よりも 1,096 日多い。これは年間平均で僅か 16 日（6%）という差異である。一方、興味深いことに、対前日の変化率の平均値は逆に下降率が上昇率を 0.0237% 上回っており、微妙なバランスが保たれていることがわかる。

表 1

	頻度	変化率
上昇	9,388	0.6503%
下降	8,292	-0.6740%
差異	1,096	0.0237%

本日の株価 x_n は前日の株価 x_{n-1} が何パーセントかの変化をして形成される。その変化率を w_n とすると、本日の株価 x_n は、

$$x_n = x_{n-1}(1 + w_n) \quad \dots \textcircled{1}$$

である。同様に、前日の株価 x_{n-1} は一昨日の株価 x_{n-2} と変化率 w_{n-1} から構成されているので、

$$x_{n-1} = x_{n-2}(1 + w_{n-1}) \cdots \textcircled{2}$$

である。②を①に代入すると、

$$x_n = x_{n-2}(1 + w_{n-1})(1 + w_n) \cdots \textcircled{3}$$

これを繰り返していくと初期値 x_0 から n 日後の株価 x_n を得る式になる。対前日比がプラスのときの変化率を a_n 、マイナスのときの変化率を b_n とすると、 a_n と b_n はランダムに表れるので、 n 日後の株価は、

$$x_n = x_0(1 + a_1)(1 - b_1)(1 + a_2)(1 + a_3)(1 - b_2)(1 - b_3) \cdots \textcircled{4}$$

$(a, b > 0)$

のように書ける。変化率がプラスであった日の合計を p 、マイナスであった日の合計を q だとすると、④は、

$$x_n = x_0 \{(1 + a_1)(1 + a_2)(1 + a_3) \cdots (1 + a_p)\} \{(1 - b_1)(1 - b_2)(1 - b_3) \cdots (1 + b_q)\} \cdots \textcircled{5}$$

と書き換えることができる。変化率の掛け算が長々と続くので、単純に変化率の平均値に置き換えてみよう。そうすると⑤式は、

$$x_n = x_0 \left(1 + \frac{a_1 + a_2 + a_3 + \cdots + a_p}{p}\right)^p \left(1 - \frac{b_1 + b_2 + b_3 + \cdots + b_q}{q}\right)^q \cdots \textcircled{6}$$

すなわち、

$$x_n = x_0 (1 + \bar{a}_p)^p (1 - \bar{b}_q)^q \cdots \textcircled{7}$$

$(p + q = n, a, b > 0)$

となり、すっきりした形になる。⑦式は正負の変化率の平均値とそれぞれの日数が分かれば n 日後の株価が計算できるということを示している。これを S&P500 に適用してみると初期値 $x_0=16.66$ 、正負の変化率と発生日数は表 1 のとおりなので、それぞれの値を⑦式に代入すると、

$$x_{16780} = 16.66(1 + 0.6503\%)^{9388}(1 - 0.6740\%)^{8292} = 1974.86$$

となる。17,680 日後の実際の株価は 2799.55 であるので 800 ドル以上低い値である。ではこの簡易的な計算方法に使い道はないのだろうか。ここで期間を 70 年から 1 年に短縮して

みよう。表 1 が直近の 10 年間の実績と計算結果の比較である。右端の差異というのが計算式⑦で得られた値と実際の値（各年度末の終値）の差である。平均すると 0.4%の差異にとどまっており、この程度であれば予測に使いそうである。

表 1

	平均上昇率	平均下降率	上昇日数	下降日数	取引日数	計算値 x_n	実績	差異
2010	0.747%	-0.869%	144	108	252	1267.97	1,257.64	0.814%
2011	0.957%	-1.135%	138	114	252	1274.59	1,257.60	1.333%
2012	0.611%	-0.570%	132	118	250	1431.48	1,426.19	0.370%
2013	0.554%	-0.522%	147	105	252	1853.09	1,848.36	0.255%
2014	0.510%	-0.560%	143	109	252	2024.23	2,023.03	1.684%
2015	0.747%	-0.686%	121	132	253	2041.07	2,043.94	0.140%
2016	0.598%	-0.555%	130	122	252	2,248.61	2,238.83	-0.029%
2017	0.327%	-0.269%	143	108	251	2,669.47	2,673.61	-0.155%
2018	0.680%	-0.822%	134	117	251	2,546.60	2,506.85	1.561%
2019	0.580%	-0.576%	148	104	252	3,158.90	3,221.29	-1.975%
平均	0.631%	-0.656%	138	114	252	2,051.60		0.400%

2) 上昇・下降日数と変化率のパターン分析と近未来予測

繰り返しになるが、株価は日数と変化率の組み合わせにより決定されるわけだから、組み合わせとして以下の四つのが考えられる（表 2）。組み合わせの A は上昇した日数の方が下降した日数よりも多く、上昇率も下降率より高かった年。B は上昇日数の方が下降日数よりも多かったが、上昇率が下降率より低かった年。C は上昇日数が下降日数よりも少なかったが、上昇率が下降率よりも高かった年。D は上昇日数と上昇率が下降日数と下降率よりも低かった年である。発生確率が最も高いのがパターン A で 38.6%、二番目に高いのがパターン B で 34.3%、以下、パターン C の 14.3%、D の 12.9%と続く。表 1 の色分けは表 2 に対応している。

表 2

変化パターン		計	発生確率	発生確率
A	上昇日数 > 下降日数、上昇率 > 下降率	27	38.6%	72.9%
B	上昇日数 > 下降日数、上昇率 < 下降率	24	34.3%	
C	上昇日数 < 下降日数、上昇率 > 下降率	10	14.3%	27.1%
D	上昇日数 < 下降日数、上昇率 < 下降率	9	12.9%	
		70	100.0%	100.0%

では次に、それぞれのパターンにおける実際の上昇・下降日数、変化率をみてみよう（表 3）。

パターン A は表 2 から過去 27 回出現しているが、表 3 はさらにパターン A の平均上昇日数は 129 日、同下降日数は 104 日、平均上昇率は 0.536%、同下降率は 0.499 だっことを示している。

表 3

	平均上昇率	平均下降率	上昇日数	下降日数	取引日数
A	0.536%	-0.499%	129	104	232
B	0.583%	-0.648%	134	108	242
C	0.953%	-0.844%	129	149	278
D	0.713%	-0.785%	121	131	252
平均	0.696%	-0.694%	128	123	251

次に知りたいことは、それぞれのパターンがどのくらいの確率で出現しているか、さらにはどのパターンのあとにどのパターンが出現しやすいか、という確率である。パターンは A から D の四通りあるので、出現パターンは全部で 16 通りある。そしてその 16 通りが実際どの程度の確率で発生したかを示したものが表 4 である。

表 4

	データ数	発生確率	データ数	発生確率	発生確率
A→A	10	14.3%	26	37.1%	70.0%
A→B	9	12.9%			
A→C	3	4.3%			
A→D	4	5.7%			
B→A	7	10.0%	23	32.9%	
B→B	10	14.3%			
B→C	3	4.3%			
B→D	3	4.3%			
C→A	7	10.0%	12	17.1%	30.0%
C→B	2	2.9%			
C→C	1	1.4%			
C→D	2	2.9%			
D→A	4	5.7%	9	12.9%	
D→B	1	1.4%			
D→C	4	5.7%			
D→D	0	0.0%			
	70	100.0%	70	100.0%	100.0%

表5はパターン別に翌年出現するパターンの確率を示したものである。たとえば2020年はパターンAであったので翌年もパターンAである確率は38.5%ある一方で、パターンDになる可能性も15.4%あるということを示している。過去データが合計70しかないのでこの確率はかなりおおざっぱの域は出ないが、目安にはなろう。

表5

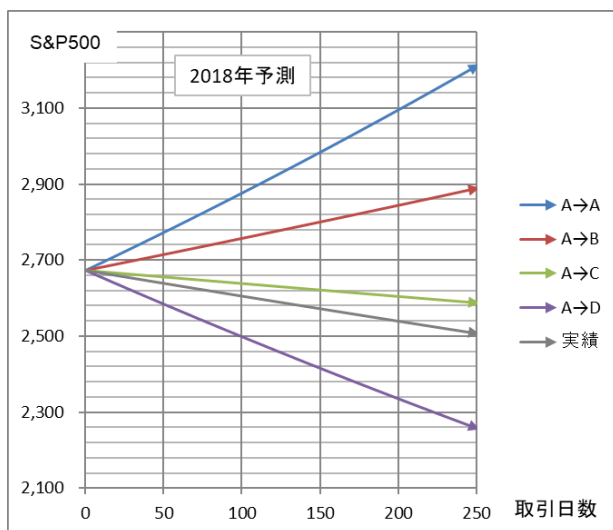
	データ数	計	発生確率
A→A	10	26	38.5%
A→B	9		34.6%
A→C	3		11.5%
A→D	4		15.4%
B→A	7	23	30.4%
B→B	10		43.5%
B→C	3		13.0%
B→D	3		13.0%
C→A	7	12	58.3%
C→B	2		16.7%
C→C	1		8.3%
C→D	2		16.7%
D→A	4	9	44.4%
D→B	1		11.1%
D→C	4		44.4%
D→D	0		0.0%

さて、ここまでくれば計算式⑦を使って、翌年の株価を確率的に予測することができる。初期値は昨年度の終値、パターン別の上昇・下降日数、変化率の値は表3、次に現れるであろうパターンの確率は表5をみればわかる。表6と図1が2017年の終値から2018年の終値を予測したとき、表7と図2が2018年から2019年を予測したときの予測値と実績である。

表6

2018年予測	出現実績	確率	終値予測	増減率
A→A	10	40.0%	3,214.57	20.2%
A→B	8	32.0%	2,891.63	8.2%
A→C	3	12.0%	2,588.02	-3.2%
A→D	4	16.0%	2,255.44	-15.6%
	25	100.0%		

図 1

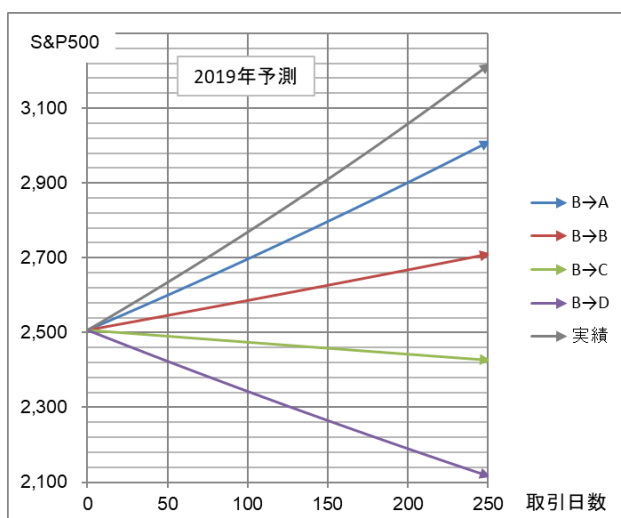


2017年の終値=2673.61に対して、2018年の終値は2506.85だった。

表 7

2019年予測	出現実績	確率	終値予測	増減率
B→A	6	27.3%	3,014.07	20.2%
B→B	10	45.5%	2,711.27	8.2%
B→C	3	13.6%	2,426.60	-3.2%
B→D	3	13.6%	2,114.76	-15.6%
	22	100.0%		

図 2



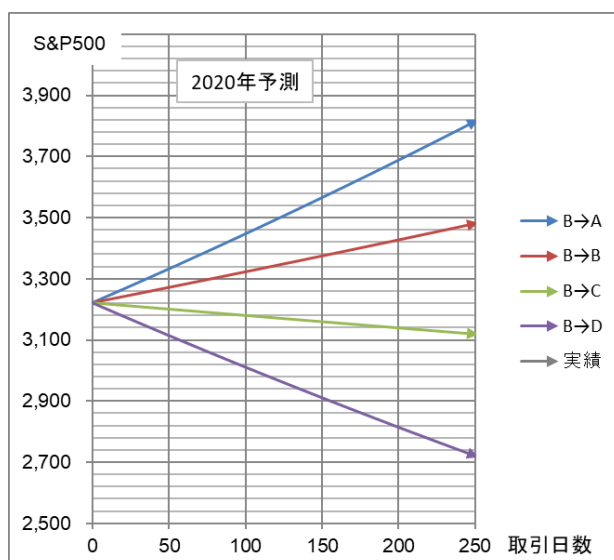
2018年の終値=2506.85に対して2019年の終値は3221.29だった。

それでは最後に今年度（2020年）の株価を予想してみよう。2019年の終値 3221.29 ドルに対して、2020年の終値の予測幅は 2,771 ドルから 3,822 ドルである。

表 8

2020年予測	出現実績	確率	終値予測	増減率
A→A	10	38.5%	3,821.99	18.6%
A→B	9	34.6%	3,483.97	8.2%
A→C	3	11.5%	3,118.16	-3.2%
A→D	4	15.4%	2,717.46	-15.6%
	26	100.0%		

図 3



2019年の終値=3221.29に対して、2020年の終値は？

3) 結論

株価が形成される単純な四つの要素に注目し、予測値を算出する計算式を導き、その計算値と実績値を比較することによって一年くらいのスパンであれば計算式から近似値が得られることを確認した。さらに、過去70年間において四つの要素がどのように組み合わせたり変移したかをパターン分析することによって、次年度の株価が到達するであろう領域を確率的に予測できることを示した。非常に単純な手法であるが、具体的な予測株価を計算できるところがこの手法の特長である。次の頁に過去70年間のデータを示したので参考にされたい。

		平均上昇率	平均下降率	上昇日数	下降日数	取引日数
1	1950	0.667%	-0.733%	140	108	248
2	1951	0.565%	-0.520%	130	119	249
3	1952	0.392%	-0.388%	135	115	250
4	1953	0.411%	-0.486%	127	124	251
5	1954	0.515%	-0.457%	153	100	253
6	1955	0.626%	-0.736%	153	99	252
7	1956	0.680%	-0.612%	115	136	251
8	1957	0.543%	-0.618%	120	132	252
9	1958	0.479%	-0.457%	157	95	252
10	1959	0.440%	-0.452%	136	116	252
11	1960	0.498%	-0.508%	123	130	253
12	1961	0.400%	-0.384%	148	102	250
13	1962	0.685%	-0.718%	120	132	252
14	1963	0.397%	-0.376%	144	107	251
15	1964	0.281%	-0.286%	147	106	253
16	1965	0.310%	-0.326%	141	111	252
17	1966	0.525%	-0.587%	120	132	252
18	1967	0.416%	-0.397%	144	107	251
19	1968	0.434%	-0.426%	121	105	226
20	1969	0.474%	-0.550%	122	128	250
21	1970	0.699%	-0.689%	126	128	254
22	1971	0.475%	-0.485%	138	115	253
23	1972	0.421%	-0.387%	138	113	251
24	1973	0.794%	-0.774%	113	139	252
25	1974	1.126%	-1.035%	105	148	253
26	1975	0.850%	-0.736%	135	118	253
27	1976	0.611%	-0.545%	134	119	253
28	1977	0.435%	-0.478%	118	134	252
29	1978	0.604%	-0.623%	129	123	252
30	1979	0.500%	-0.536%	140	113	253
31	1980	0.826%	-0.823%	141	112	253
32	1981	0.652%	-0.668%	121	132	253
33	1982	1.014%	-0.721%	114	139	253
34	1983	0.663%	-0.655%	138	115	253
35	1984	0.692%	-0.542%	113	140	253
36	1985	0.550%	-0.452%	137	116	253
37	1986	0.652%	-0.690%	141	112	253
38	1987	1.022%	-1.260%	143	110	253
39	1988	0.724%	-0.768%	138	115	253
40	1989	0.572%	-0.596%	150	102	252
41	1990	0.704%	-0.852%	135	118	253
42	1991	0.779%	-0.559%	124	129	253
43	1992	0.469%	-0.456%	130	124	254
44	1993	0.416%	-0.381%	130	123	253
45	1994	0.435%	-0.495%	133	119	252
46	1995	0.398%	-0.337%	156	96	252
47	1996	0.586%	-0.531%	138	116	254
48	1997	0.870%	-0.831%	140	113	253
49	1998	0.916%	-0.916%	140	112	252
50	1999	0.960%	-0.849%	129	123	252
51	2000	1.083%	-1.047%	120	132	252
52	2001	1.019%	-1.031%	119	129	248
53	2002	1.317%	-1.219%	112	140	252
54	2003	0.855%	-0.803%	137	115	252
55	2004	0.522%	-0.570%	140	112	252
56	2005	0.473%	-0.570%	141	111	252
57	2006	0.472%	-0.476%	140	111	251
58	2007	0.675%	-0.770%	137	114	251
59	2008	1.585%	-1.904%	126	127	253
60	2009	1.201%	-1.280%	140	112	252
61	2010	0.747%	-0.869%	144	108	252
62	2011	0.957%	-1.135%	138	114	252
63	2012	0.611%	-0.570%	132	118	250
64	2013	0.554%	-0.522%	147	105	252
65	2014	0.510%	-0.560%	143	109	252
66	2015	0.747%	-0.686%	121	132	253
67	2016	0.598%	-0.555%	130	122	252
68	2017	0.327%	-0.269%	143	108	251
69	2018	0.680%	-0.822%	134	117	251
70	2019	0.580%	-0.576%	148	104	252