Introduction

National governments often spend massively in poor areas to try to lift them out of poverty permanently. Examples include the U.S. War on Poverty in Appalachia in the 1960s as well as federal spending on U.S. possessions in Micronesia. Such programs presume that a few, identifiable causes underlie area poverty. Suppose, however, that area income, relative to national income, follows a random walk. Then the relative area income is unlikely to be due to a few large factors. A targeted anti-poverty program may be unlikely to succeed. On the other hand, if area income follows a random walk only after the program has taken effect, then that may provide evidence of the program's success; for it may suggest that the program has removed systematic causes of area poverty, leaving only unsystematic factors to influence it.

The goal of this paper is to determine whether relative income in Appalachia followed a random walk before the War on Poverty.

Section II describes the analysis. It treats a random walk as a pattern in which a variable, such as relative income, is as likely to rise as to fall in any time period of a given length. One may test, for randomness, the pattern of accumulated rises and falls that actually occurred in relative income. The analysis focuses on the personal income per capita of the one state to fall entirely within Appalachia, West Virginia, relative to the counterpart income of the nation. Section III presents results. Section IV presents conclusions and reflections.

Analysis

The paper's concept of a random walk follows that of Feller (ref: feller , chapter 3). In a random walk, a variable X_t is equally likely to lead to outcome A or to outcome B in repeated experiments, where t indexes the repetitions. Either A or B must occur. If experiment t results in A, then let $X_t = 1$; if B results instead, then let $X_t = -1$. To easily track the experiments as they unfold, let $S_t = \sum_{i=0}^{t} X_i$ and track this partial sum over t. The paper will define the path of S_t over $t = 0, 1, 2, ..., n$ as a *walk with n epochs*.

Testable characteristics of random walk

The partial sum of a random walk follows probabilistic laws. In deciding whether a time series follows a random walk, one may study whether its characteristics are improbable for a random walk. For example, one may construct a partial sum from the time series. Then one may study two features that summarize much of the behavior of the partial sum in a random walk: Its sign and its extrema. For those features, four characteristics in particular may disclose whether a series is random:

(1) *How often does it change sign*? More frequent sign changes may indicate that the walk is more likely to be random. But random walks often rarely change sign. Given that a walk of fixed length is random, the probability that it will change sign $k = 1, 2, ..., n$ times decreases as k increases. While a random walk may be more likely to change sign often than a nonrandom walk, additional tests for randomness are still required.

(2) *When did the walk touch or cross the horizontal axis for the last time*? Given the number of sign changes, a random walk may be more likely to touch or cross the horizontal axis than a nonrandom walk. If the path touches the horizontal axis in epoch m, then denote the point as $(m, 0)$ and the event as a *visit to the origin*. Denote the probability of the event as u_m.

Passing both tests (1) and (2) is not sufficient to sustain a conclusion of randomness. For example, the pattern in a walk may call for it to change sign often at the end. Additional tests are required.

(3) *What are the extrema of the walk*? Given the number of sign changes and the length of time since the last visit to the horizontal axis, a random walk may have a lower maximum or a higher minimum than a nonrandom walk.

Tests (1), (2) and (3) do not comprise a sufficient battery for randomness. For example, the pattern in a walk may be to achieve a high maximum or a low minimum and then to change sign often at the end. The battery requires a fourth test.

(4) *When does the walk first reach a maximum or minimum*? Given the number of sign changes, the time of the last visit to the horizontal axis, and the levels of the extrema, a random walk may attain an extremum later than a nonrandom walk.

Algorithms for the four tests follow. Feller (cite: feller , Chapter 3) provides proofs related to the probabilities used here.

Tests for randomness

Test 1: *Number of sign changes*

One expects a random walk to change sign more often than a nonrandom walk. Suppose that the walk under consideration makes R sign changes up through epoch $2n+1$. (A sign change may occur only in an odd-numbered epoch, since the number of steps of one sign must exceed the number of steps of the other sign by one.) One might then test a walk for randomness by computing the probability that a random walk would change sign R times or fewer. If the probability is low, then one might say that the walk has failed that test for randomness.

In particular, the probability that a random walk would change sign exactly R times up through epoch $2n+1$ is twice the probability that it would achieve level $2R+1$ in epoch $2n+1$. The paper will denote this equivalent of the probability of R sign changes as $2P\{S_{2n+1} = 2R+1\}$. If, for the walk under consideration, it is true that

$$\sum_{i=0}^{R} 2P\{S_{2n+1} = 2R+1\} < .01, \qquad \#$$

then the analysis will conclude that the walk has failed the sign-change test for randomness. If (ref: sign-change test) holds true, then the walk will have passed the test.

Now consider the general probability that the partial sum of a random walk attains value k in epoch m. (For concision, the paper may simply say that the walk attains value k in epoch m). The paper will denote the probability as $P\{S_m = k\}$ or as $p_{m,k}$. The number of ways in which the partial sum of a random walk may attain value k in epoch m is

$$\binom{m}{\frac{m+k}{2}}.$$

If the binomial is positive, then k and m have the same parity; if they differ in parity, then the binomial is zero.

One important special case might illustrate the principle behind the formula: The number of ways that the walk may visit the origin in epoch m. For that event, half the epochs must have contained positive steps. The number of ways that the event can occur, then, is the number of ways that half of the m epochs can be chosen to contain positive steps:

$$\binom{m}{\frac{m}{2}}.$$

Return now to the general probability that a random walk leads up to point (m, k). The total number of paths feasible for a random walk with m epochs is 2^m, since the walk may take either of two branches at any epoch. Every feasible path is equally probable for a random walk. So, the probability that the walk would reach point (m, k) is the share, in all feasible paths, of the paths that include (m, k):

$$P\{S_m = k\} = \binom{m}{\frac{m+k}{2}} 2^{-m}. \qquad \#$$

Applying this formula to (ref: sign-change test) yields

$$2^{-2n} \sum_{i=0}^{R} \binom{2n+1}{i+n+1} < .01.$$

Test 2: *Time of the last visit to the origin*

From the perspective of the last epoch, a random walk is likely to have visited the origin more recently than a nonrandom walk, since positive and negative steps are more likely to offset one another in a random walk than in a nonrandom walk. Suppose that the last visit to the origin for a walk occurs in epoch $2k$. Then one test for randomness could compute the probability that a random walk of the same length would have last visited the origin in epoch $2k$ or earlier. If this probability is small, then one may say that the walk has failed the last-visit test for randomness.

Consider now the probability of a last visit for a random walk. Any visit to the origin can occur only in an even-numbered epoch. The probability that the last visit occurs in epoch $2k$, in a walk with $2n$ epochs in all, is $u_{2k}u_{2n-2k}$.

Consider a walk that has $2n$ epochs in all and that has a last visit in epoch $2k$. If, for that walk, it is true that

$$\sum_{i=0}^{k} u_{2i}u_{2n-2i} < .01, \qquad \#$$

then the analysis will conclude that the walk has failed the last-visit test for randomness. If (ref: last visit) does not hold true, then the walk has passed the test.

Applying the formula in (ref: Formula) to (ref: last visit) yields an algorithm for computation:

$$\frac{1}{2^{2n}} \sum_{i=0}^{k} \binom{2i}{i} \binom{2n-2i}{n-i} < .01.$$

Test 3: *Extrema of the walk*

The partial sums of a random walk are likely to have a maximum or a minimum that is smaller in absolute value than a nonrandom walk, since positive steps are more likely to offset negative steps in a random walk than in a nonrandom walk. Suppose that the partial sums in a walk have a maximum value of v. Then one may test for randomness in the walk by computing the probability that the partial sums in a random walk of the same length would have a maximum of v or greater. If the probability is small, then one may conclude that the walk has failed the extrema test for randomness.

The probability that the partial sum in a random walk of n epochs has a maximum of v equals $p_{n,v} + p_{n,v+1}$. Since the number of the epoch and the value of the partial sum in that epoch must have the same parity, $p_{n,v} > 0$ implies that $p_{n,v+1} = 0$; and $p_{n,v+1} > 0$ implies that $p_{n,v} = 0$.

If, for a walk with n epochs and maximum v, it is true that

$$\sum_{i=v}^{n} p_{n,i} + p_{n,i+1} < .01. \qquad \#$$

then the analysis will conclude that the walk has failed the extrema test for randomness. If (ref: extrema test) does not hold true, then the walk has passed the test.

Applying (ref: Formula) to (ref: extrema test) yields an algorithm for computation:

$$\frac{1}{2^n} \sum_{i=v}^{n} \left[\binom{n}{\frac{n+i}{2}} + \binom{n}{\frac{n+1+i}{2}} \right] < .01. \qquad \#$$

Since the numbers of epochs are integers, the binomial with a noninteger in its denomination should be set to zero. Computations may thus be aided by this re-expression of (ref: actual extrema test):

$$\frac{1}{2^n} \left[\sum_{j=\frac{v}{2}}^{\frac{n}{2}} \binom{n}{\frac{n+2j}{2}} + \sum_{j=\frac{v}{2}}^{\frac{n}{2}-1} \binom{n}{\frac{n+1+2j+1}{2}} \right] < .01$$

when v is even; or

$$\frac{1}{2^n} \left[\sum_{j=v+1}^{\frac{n}{2}} \binom{n}{\frac{n+2j}{2}} + \sum_{j=v}^{\frac{n}{2}-1} \binom{n}{\frac{n+1+2j+1}{2}} \right] < .01$$

when v is odd. Both expressions assume that n is even. footnote

The test in (ref: actual extrema test) also holds for a minimum of $-v$. To test whether a walk with n epochs and a partial-sum minimum of $-v$ is random, use the absolute value of $-v$ in (ref: actual extrema test).

Whether to apply the maximum test or the minimum test for randomness depends on whether the partial sums in the random walk are mainly positive or negative. In a walk of n epochs, if $S_i \geq 0$ for at least $n/2$ values of i, the index for epochs, then apply the maximum test; otherwise, apply the minimum test.

Test 4: *First extremum*

Positive and negative steps are more likely to offset one another in a random walk than in a nonrandom walk. So, one would expect that a random walk would first reach a given maximum in an epoch later than the one in which a nonrandom walk of the same length would first reach the same given maximum. Consider a walk with n epochs (where, for convenience, n is taken to be an even number) that first attains its maximum v in epoch k. One may test the walk for randomness by computing the probability that a random walk of length n would first reach its maximum v at point (k, v) or in an epoch earlier than k. If the probability is small, then one may say that the walk has failed the first-extremum test for randomness.

The probability that a random walk, with n epochs in all, would first achieve its maximum in epoch k is $\frac{1}{2}u_{2p}u_{n-2p}$, where $k = 2p$ (if k is even) or $k = 2p + 1$ (if k is odd). Consider a walk that has n epochs and that first reaches its maximum in epoch k. Suppose, for that walk, that it is true that

$$\frac{1}{2}\sum_{i=0}^{k/2} u_{2i}u_{n-2i} < .01 \qquad \#$$

when k is even; or that

$$\frac{1}{2}\sum_{i=0}^{(k-1)/2} u_{2i}u_{n-2i} < .01 \qquad \#$$

when k is odd. Then the analysis will conclude that the walk has failed the first-extremum test for randomness. If (ref: first-extremum test one) holds for even k, or (ref: first-extremum test two) holds for odd k, then the analysis will conclude that the walk has passed the test.

The formulas in (ref: first-extremum test one) and (ref: first-extremum test two) also hold for a minimum that first occurs in epoch k. If the maximum test was applied in (3), then apply now the first-maximum test. If the minimum test was applied in (3), apply now the first-minimum test.

Applying (ref: Formula) to (ref: first-extremum test one) and to (ref: first-extremum test two) yields these algorithms for computation:

$$\frac{1}{2^{n+1}}\sum_{i=0}^{k/2}\left[\binom{2i}{i}\binom{n-2i}{\frac{n}{2}-i}\right] < .01$$

if k is even; or

$$\frac{1}{2^{n+1}}\sum_{i=0}^{(k-1)/2}\left[\binom{2i}{i}\binom{n-2i}{\frac{n}{2}-i}\right] < .01$$

if k is odd.

Summary of tests

The analysis will conclude that the walk is random only if it passes all four tests for randomness. If the walk fails one or more of the four tests for randomness, then it will be characterized as nonrandom.

The probability that a random walk would fail one or more of the four tests is low:

$$\sum_{i=1}^{4}\binom{4}{i}(.01)^{i}(.99)^{4-i} < .04.$$

The test standard for randomness seems fairly stringent. I have selected it because a conclusion that relative income follows a random walk, when it actually does not, may carry more harmful policy implications than a conclusion that relative income does not follow a random walk when it actually does.

Implementing the tests

For the study period of n years, n is chosen for convenience to be an even number. Test 1 remains valid since the probability of a sign change in an even-numbered epoch is zero; the length of the walk for that test is taken to be $n + 1$. The analysis denotes the year as t, where $0 \leq t \leq n$. Let $X_t = 1$ if relative income rose that year over the year before; otherwise, $X_t = -1$. Let the partial sum of the walk in epoch i be

$$S_i = \sum_{t=1}^{i} X_t.$$

The table below gives data for the tests. Annual personal income per capita for West Virginia is

WVY (Series F346, cite: census). Annual personal income per capita for the United States is *USY* (Series F297, cite: census). The ratio of West Virginian personal income to U.S. personal income, per capita, is *WVY*/*USY*. The average annual ratio over the period from 1948 through 1970 is .74. The ratio series provides the basis for estimation of the walk. The epoch is indexed by t, where 1948 corresponds to epoch 0. The end year 1970 is chosen to truncate the time series before permanent economic effects of the War on Poverty in the region would become evident.

Epoch (t)	**Year**	**WVY**	**USY**	**WVY/USY**	**X**	**S**
0	1948	1120	1430	0.7832		0
1	1949	1033	1384	0.7464	-1	-1
2	1950	1065	1496	0.7119	-1	-2
3	1951	1192	1652	0.7215	1	-1
4	1952	1258	1733	0.7259	1	0
5	1953	1282	1804	0.7106	-1	-1
6	1954	1232	1785	0.6902	-1	-2
7	1955	1326	1876	0.7068	1	-1
8	1956	1491	1975	0.7549	1	0
9	1957	1610	2045	0.7873	1	1
10	1958	1565	2068	0.7568	-1	0
11	1959	1600	2161	0.7404	-1	-1
12	1960	1612	2216	0.7274	-1	-2
13	1961	1658	2265	0.7320	1	-1
14	1962	1727	2370	0.7287	-1	-2
15	1963	1819	2458	0.7400	1	-1
16	1964	1943	2590	0.7502	1	0
17	1965	2087	2770	0.7534	1	1
18	1966	2250	2987	0.7533	-1	0
19	1967	2403	3170	0.7580	1	1
20	1968	2545	3436	0.7407	-1	0
21	1969	2738	3708	0.7384	-1	-1
22	1970	3047	3943	0.7728	1	0

1

State and national incomes per capita tended to rise and fall together, but relative income may have cycled mildly every 10 years, with peaks in 1957 and 1967. The partial sums S_t, follow a walk that may look more random than nonrandom to the eye. Extrema are modest and returns to the origin are frequent. Quantitative tests, however, may permit a more systematic evaluation of any randomness. They are discussed below.

Test 1: Sign changes

There are four sign changes in S_t, in the years 1957, 1959, 1965 and 1969. The probability that a random walk would change signs four times or fewer in 21 epochs is

$$2^{-20}\sum_{i=0}^{4}\binom{21}{i+11} = .97$$

so that the walk passes this test for randomness.

Test 2: Last visit to origin

The walk last visits the horizontal axis in the last epoch, 22. The probability that a random walk would last visit the origin in the last epoch or earlier (including epoch 0) is 1. For purposes of illustration, the calculation for a walk of 22 epochs is:

$$\frac{1}{2^{22}}\sum_{i=0}^{11}\binom{2i}{i}\binom{22-2i}{11-i} = 1.$$

The walk passes the last-visit test for randomness.

Test 3: Minimum test

Since more than half of the partial sums are negative – 12 of 22 – the extremum test focuses on the minimum (not the maximum) of the walk. The minimum is -2. The probability that a random walk of 22 epochs would have a minimum of -2 or greater in absolute value is

$$\frac{1}{2^{22}}\left[\sum_{j=1}^{11}\binom{22}{\frac{22+2j}{2}}+\sum_{j=1}^{10}\binom{22}{\frac{24+2j}{2}}\right] = .68$$

so that the walk passes the minimum test for randomness.

Test 4: First minimum

The walk first reaches its minimum, -2, in epoch 2. The probability that a random walk of 22 epochs would first achieve its minimum of -2 in epoch 2 or earlier is

$$\frac{1}{2^{23}}\sum_{i=0}^{1}\left[\binom{2i}{i}\binom{22-2i}{11-i}\right] = .13.$$

The walk passes the first-minimum test but not by the substantial margin of the other tests.

Conclusions and reflections

The tests in this paper do not suggest to me that personal income per capita in West Virginia, relative to that of the nation, from 1948 through 1970, was determined by a few large factors. The paper has applied four tests for randomness to that income trend. For none of the four tests was I able to reject the hypothesis that the trend was, indeed, random. For three of four tests, the probability that a random walk would produce the same characteristic as that observed in the income trend exceeded 67 percent. For the fourth test, the probability was but 13 percent; whether one regards this as evidence of randomness may be a more subjective judgment than for the preceding three tests.

These results need not imply that the War on Poverty in Appalachia wasted money. That conclusion may hinge on one’s definition of randomness. The goal of this paper instead has been to raise this question: How might an organization determine the likelihood of success for a program that is aimed at solving a problem by ameliorating a few assumed causes?

feller Feller, William. *An introduction to probability theory and its applications*. Third edition, revised printing. Volume I. New York: John Wiley & Sons, 1970.

census U. S. Bureau of the Census. *Historical statistics of the United States, Colonial times to* 1970. Part I. Washington, D.C., 1975.