
# DEATHDAY AND BIRTHDAY: An Unexpected Connection

**David P. Phillips** *State University of New York, Stony Brook*

In the movies and in certain kinds of romantic literature, we sometimes come across a deathbed scene in which a dying person holds onto life until some special event has occurred. For example, a mother might stave off death until her long-absent son returns from the wars. Do such feats of will occur in real life as well as in fiction? If some people really do postpone death, how much can the timing of death be influenced by psychological, social, or other identifiable factors? Can deaths from certain diseases be postponed longer than deaths from other diseases?

In this essay we shall see how dying people react to one special event: their birthdays. We want to learn whether some people postpone their deaths until after their birthdays. If we compare the date of death with the date of birth for a large number of people, will we find fewer deaths than expected just before the birthday? If we do find a dip in deaths, we may conclude that some of these people are postponing death until after their birthdays.

We shall use elementary statistical methods in approaching the problem. For example, the comparison of an actual number of events with the number that might be expected is one of these methods; others will be noted later.

## NATURE OF THE DATA TO BE INVESTIGATED

We shall examine only the deaths of famous people. There are two reasons for this. First, it seems likely that ordinary people look forward to their birthdays less eagerly than do famous people because a very famous person's birthday, generally, is celebrated publicly, and he may receive a substantial amount of attention, gifts, and so on. In contrast, much less attention is paid to the birthday of an ordinary person, and he may have relatively little reason to look forward to it. Hence famous people may be more likely to postpone deaths than less famous ones. Second, it is easier to examine the deaths of the famous than of other people. To discover whether there is a dip in deaths before the birthday, we need information on the birth and death dates of individuals. This type of information is not available from conventional tables of vital statistics; therefore, we cannot easily determine the birth and death dates of large numbers of ordinary people. On the other hand, we can easily determine the birth and death dates of famous people because there is much biographical information available about them.

In all, we shall examine the birth and death dates of more than 1200 people. It is tedious to classify these dates by *day,* so we shall examine the *month* of birth and *month* of death. Thus, for the purpose of this analysis, we shall be concerned, not with the relationship between the birthday and the day of death, but rather with the relationship between the birth month and the death month. For our purposes, a person is said to have died in his birth month if the month of his death has the same name as the month of his birth. For example, if a person was born on March 1, 1897, and died on March 31, 1950, he died in his birth month. On the other hand, if he was born on March 1 and died on February 28, he did not die in his birth month; rather, he died in the month just before his birth month. Although we gain convenience by examining events by month rather than by day, we lose precision; if we find a dip in deaths in the month before the birth month, we cannot tell whether a dying person is hanging on for a few days or for a few weeks.

## IS THERE A DIP IN DEATHS BEFORE THE BIRTH MONTH?

Table 1 shows the month of birth and month of death of people listed in *Four Hundred Notable Americans.* For example, we can see from the first column that one person who was born in January died in January, two people who were born in January died in February, and so on. The column labeled "Row Total" gives the total number of people who died in each month and the row labeled "Column Total" gives the total number of people born in each month.



Table 1 enables us to compare two hypotheses. The first hypothesis states that the death month is related to (is dependent on) the birth month in that some people postpone death in order to witness their birthdays. This will be called the *death-dip hypothesis.* The second hypothesis states that no deaths are being postponed and that the month of death is not related to (is independent of) the month of birth. This will be called the *independence hypothesis.* (When we formulate our problem in terms of two hypotheses, one that we wish to disprove in order to lend credence to the other, and try to decide which hypothesis seems more consistent with the data, we are using a standard statistical testing procedure. The concept of "independence" is also an important part of many statistical hypotheses.)

Our general plan is to see whether the month immediately preceding the birth month has fewer deaths than the independence hypothesis suggests. As we explain below, it turns out that if the independence hypothesis were true, about $1/12$ of the deaths would occur in each of the six months preceding the birth month, $1/12$ in the birth month, and $1/12$ in each of the five following months. Although this may seem obvious, it actually depends upon detailed calculations because we must take into account that some calendar months produce more deaths than others and some produce more births than others. It is intuitively satisfying, nevertheless, that for the independence hypothesis, the calculations give nearly equal expected numbers of deaths for each of the twelve months preceding, during, and following the birth month.

We now compare the actual number of deaths before the birth month with the number of deaths that are expected, on the average, if the independence hypothesis is true. If the observed number of deaths is noticeably less than the expected number, there is a dip in deaths before the birth month.

First, we count the actual number of deaths in the month just before the birth month. If we sum the numbers in the starred cells in Table 1, we will have the total observed number of deaths in this period. This number is 16.

Now we calculate the total expected number of deaths in the month just before the birth month. If the independence hypothesis is true, then the death month is independent of the birth month. This means that the deaths of those born in any given month should be distributed throughout the year in the same way as the deaths of those born in any other month. Thus, because 6.32% [this is $(22/348) \times 100$] of all the deaths in Table 1 fall in December, 6.32% of those born in January should die in December, 6.32% of those born in February should die in December, and so on. In Table 1, we see that 11.2% [$39/348) \times 100$] of all deaths fall in April. Then,

TABLE 1. Number of Deaths by Month of Birth and Month of Death (Sample 1)

| MONTH OF DEATH | MONTH OF BIRTH: Jan. | Feb. | Mar. | Apr. | May | June | July | Aug. | Sept. | Oct. | Nov. | Dec. | ROW TOTAL |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Jan. | 1 | 1* | 2 | 1 | 2 | 2 | 4 | 3 | 1 | 4 | 2 | 4 | 27 |
| Feb. | 2 | 3 | 1* | 3 | 1 | 0 | 2 | 1 | 2 | 2 | 6 | 4 | 27 |
| Mar. | 5 | 6 | 5 | 3* | 1 | 0 | 5 | 1 | 2 | 5 | 3 | 1 | 37 |
| Apr. | 7 | 6 | 3 | 2 | 1* | 3 | 3 | 1 | 3 | 2 | 4 | 4 | 39 |
| May | 4 | 4 | 2 | 2 | 1 | 2* | 4 | 1 | 3 | 2 | 1 | 5 | 31 |
| June | 4 | 0 | 4 | 5 | 1 | 1 | 1* | 2 | 1 | 2 | 4 | 0 | 25 |
| July | 4 | 0 | 3 | 4 | 3 | 3 | 4 | 1* | 6 | 4 | 2 | 5 | 39 |
| Aug. | 4 | 4 | 4 | 4 | 2 | 2 | 3 | 3 | 1* | 1 | 2 | 0 | 30 |
| Sept. | 2 | 2 | 1 | 0 | 2 | 0 | 2 | 4 | 2 | 0* | 5 | 2 | 22 |
| Oct. | 4 | 2 | 2 | 3 | 2 | 2 | 2 | 3 | 3 | 1 | 4* | 5 | 33 |
| Nov. | 0 | 2 | 0 | 2 | 1 | 1 | 0 | 3 | 3 | 3 | 1 | 0* | 16 |
| Dec. | 1* | 2 | 2 | 1 | 2 | 1 | 4 | 1 | 4 | 0 | 2 | 2 | 22 |
| COLUMN TOTAL | 38 | 32 | 29 | 30 | 19 | 17 | 34 | 24 | 31 | 26 | 36 | 32 | |
| TOTAL | | | | | | | | | | | | | 348† |

Source: R. B. Morris, ed., *Four Hundred Notable Americans* (New York: Harper & Row, 1965).
* Deaths corresponding to month preceding birth month.
† The total number of deaths is less than 400 because (1) some of those in the source volume have not yet died; (2) for some of those in the volume, the month of birth and/or death is not known.





if independence holds, we would expect 11.2% of those born in any month to die in April. For example, there are 19 people born in May; we expect 11.2% of these people, that is, $(11.2/100) \times 19 = 2.1$, to die in April. In a similar fashion, we can work out the expected number of deaths in each of the 12 starred cells in Table 1. If we sum these 12 numbers, we will have the total number of deaths that we expect to occur in the month before the birth month *if* the independence hypothesis is true. This expected number is 28.3.

The more intuitive method mentioned earlier estimates the total expected number of deaths in the starred cells to be simply $348/12 = 29.0$. In other words, we expect about $1/12$ of all deaths to occur one month before the birth month. In fact, we expect about $1/12$ of all deaths to occur in the birth month or in any month before or after it. In general, this rough-and-ready method of estimating the total expected number for any month gives results very close to those provided by more precise methods.

We can now compare the observed number of deaths before the birth month with the expected number in this period. No matter which "expected number" we use, it is considerably higher than the number of deaths observed just before the birth month: we expect about 28 or 29 deaths, but we observe only 16—about 12 fewer than expected. In other words, we observe a dip in deaths in the month before the birth month as predicted by the death-dip hypothesis. As we shall soon see, the discrepancy is much more than might reasonably be explained by chance.

## IS THERE A DEATH RISE AFTER THE BIRTH MONTH?

If the death-dip hypothesis is true, what will become of the 12 or so people who presumably have postponed death until their birthdays? When are they expected to die? There is no way of answering this question a priori because, even if the death-dip hypothesis is true, the death dip might have come about in a number of different ways; some of these different ways imply differing periods of survival for those who have postponed death. For example, the death dip could result entirely because some people who were hovering between life and death unexpectedly recovered; in this case, it might be years before these people die, and we could expect no rise in deaths immediately after the birth month. On the other hand, the death dip could appear solely because those who do not die just before their birthdays live a few days or weeks longer than expected; in this case, there should be a peak in deaths soon after the birth month. Depending on the way in which the death dip came about, we would or would not expect a rise in deaths after the birth month: we cannot tell what to expect on the basis of the death-dip hypothesis.

Although the death-dip hypothesis is not helpful here, practical experience with another sample (of famous Englishmen) suggests that we look for a rise in deaths in the four-month period consisting of the birth month and the three months thereafter. Thus, although we searched for a *death dip* in a *one-month* period, we shall search for a *death rise* in a *four-month* period, because past experience, not theory, makes this approach seem promising.

Table 2 gives the observed number of deaths six months before the birth month, five months before, and so on, down to zero months before, one month after, and so on up to five months after the birth month. Because $n = 348$ is the total number of people in sample 1, $n/12 = 29.0$ is the number expected to die six months before the birth month, five months before, and so on. From this table it is evident that not only is there a dip in deaths before the birth month, but there is also a rise in deaths during the birth month and during the three months thereafter. We expect about $4/12$ of all deaths in the first sample $[348 \times (4/12) = 116]$ to fall in this four-month period, but we observe 140 deaths during this time.

## COULD THE DEATH DIP AND DEATH RISE BE DUE TO CHANCE?

We know that surprising phenomena sometimes occur just by chance and for no other reason. For example, a person might deal himself a straight in poker; ordinarily, we attribute this happy event to the vagaries of providence, not to the dishonesty of the dealer. In much the same way, we might wonder whether the death dip and death rise have arisen by chance and for no other reason.

Now suppose our poker player were to deal himself not just one straight, but four straights in a row in the four times he deals while playing with us. This *could* have happened by chance, but it is so unlikely that we would prefer some other explanation. The less likely an event is to occur by chance, the more we prefer some other explanation. Similarly, if we find a death dip and death rise in, say, four samples and not just one, there is a small possibility that these phenomena could have occurred by chance, but another explanation might be more plausible.

In sample 1 we observed that there are fewer deaths than expected in the month before the birth month, and more deaths than expected in the four-month period consisting of the birth month and the three months thereafter. Can we find a similar death dip and death rise in other samples of people?

## MORTALITY IN THREE MORE SAMPLES

Three new samples were taken, consisting of people who are famous for two reasons. First, they achieved high status in their lifetimes: they were listed in *Who Was Who in America*. Second, they came from well-known families (e.g., Adams, Vanderbilt, Rockefeller) which are listed in "The Foremost Families of the U.S.A.," an appendix to *Royalty, Peerage and Aristocracy*

TABLE 2. Number of Deaths, Before, During, and After the Birth Month (Sample 1)

| | 6 MONTHS BEFORE | 5 MONTHS BEFORE | 4 MONTHS BEFORE | 3 MONTHS BEFORE | 2 MONTHS BEFORE | 1 MONTH BEFORE | THE BIRTH MONTH | 1 MONTH AFTER | 2 MONTHS AFTER | 3 MONTHS AFTER | 4 MONTHS AFTER | 5 MONTHS AFTER |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NUMBER OF DEATHS | 24 | 31 | 20 | 23 | 34 | 16 | 26 | 36 | 37 | 41 | 26 | 34 |

$n = 348$
$n/12 = 29.0$

Source: Table 1.

*of the World* (vol. 90, 1967). We have ensured that the new samples do not overlap each other or sample 1.

Three volumes of *Who Was Who* were examined, those for the years 1951–60, 1943–50, and 1897–1942. Sample 2 contains all who are listed in the first of these volumes and have their surnames in "Foremost Families." Sample 3 contains all who are listed in the second *Who Was Who* volume and have their surnames in "Foremost Families." Sample 4 contains *every other* person who is in the third volume of *Who Was Who* and has his surname in "Foremost Families." We chose every other person rather than every person because the third volume is so much larger than the other volumes that it would be tedious to examine every person who meets our selection criteria.

Table 3 gives the observed number of deaths before, during, and after the birth month for samples 2, 3, and 4. The last number in each row of this table is the expected number of deaths six months before the birth month, five months before, and so on. We can see that the death dip and death rise evident in sample 1 also appear in samples 2, 3, and 4. In each of these samples there are fewer deaths than expected in the month before the birth month and more deaths than expected in the four-month period consisting of the birth month and the three months thereafter.

If we now combine the data in all four samples, we get the results seen in Table 4. A graph of these results appears in Figure 1, where we can see a death dip just before the birth month and a death rise during and after it.

In summary, just before the birth month, in each of the four samples presented, we found fewer deaths than would be expected under the hypothesis that the month of death is independent of the month of birth. In each of the four samples presented, more deaths occur during and immediately after the birth month than would be expected if independence held. The similarity of results in each of the four samples helps to convince us that the death dip and death rise are real phenomena and are not merely chance fluctuations in the data.

## THE SIZE OF THE DEATH DIP AND DEATH RISE

Let us estimate the size of the death dip before the birth month and the size of the death rise thereafter for the aggregate sample of 1251 people. Out of 1251 people, 86 died in the month before the birth month. Given independence between the birth month and the death month, we would expect about $1/12$ of all 1251 deaths, or approximately 104 deaths, to fall in this period. Thus, just before the birth month only about 83% (86/104) of the deaths expected actually occurred. To put this another way, in the month before the birth month there were about 17% fewer deaths than we would expect under independence.

TABLE 3. Number of Deaths Before, During, and After the Birth Month Samples 2, 3, and 4

| NUMBER OF DEATHS | 6 MONTHS BEFORE | 5 MONTHS BEFORE | 4 MONTHS BEFORE | 3 MONTHS BEFORE | 2 MONTHS BEFORE | 1 MONTH BEFORE | THE BIRTH MONTH | 1 MONTH AFTER | 2 MONTHS AFTER | 3 MONTHS AFTER | 4 MONTHS AFTER | 5 MONTHS AFTER | TOTAL | TOTAL/12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sample 2* | 17 | 23 | 26 | 27 | 28 | 28 | 42 | 32 | 31 | 34 | 36 | 30 | 354 | 29.5 |
| Sample 3† | 10 | 14 | 12 | 11 | 8 | 12 | 15 | 15 | 15 | 13 | 20 | 13 | 158 | 13.2 |
| Sample 4‡ | 39 | 32 | 29 | 35 | 31 | 30 | 36 | 35 | 38 | 26 | 31 | 29 | 391 | 32.6 |

* Sample 2 excludes those listed in *Who Was Who in America 1951–1960* who died outside of that period and those listed in *Four Hundred Notable Americans.*

† Sample 3 excludes those listed in *Who Was Who in America 1943–1950* who died outside of that period or during World War II and those listed in *Four Hundred Notable Americans.*

‡ Sample 4 excludes those listed in *Who Was Who in America 1897–1942* who died outside of that period or during both World Wars and those listed in *Four Hundred Notable Americans.*

TABLE 4. Number of Deaths Before, During, and After the Birth Month, (All Samples Combined)

| | 6 MONTHS BEFORE | 5 MONTHS BEFORE | 4 MONTHS BEFORE | 3 MONTHS BEFORE | 2 MONTHS BEFORE | 1 MONTH BEFORE | THE BIRTH MONTH | 1 MONTH AFTER | 2 MONTHS AFTER | 3 MONTHS AFTER | 4 MONTHS AFTER | 5 MONTHS AFTER |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NUMBER OF DEATHS | 90 | 100 | 87 | 96 | 101 | 86 | 119 | 118 | 121 | 114 | 113 | 106 |

$n = 1251$
$n/12 = 104.3$





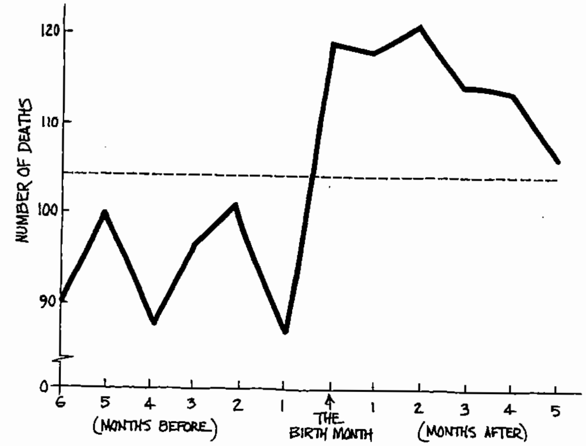


FIGURE 1
*Number of deaths before, during, and after birth month (all samples combined)*

Similarly, we can estimate the size of the death rise during the birth month and the three months thereafter. There were 472 deaths in the period including the birth month and one, two, and three months thereafter. Given independence between birth and death months, the expected number of deaths in this four-month period is estimated to be $4/12$ of all deaths, or 417 deaths $[(4/12) \times 1251 = 417]$. Thus the actual number of deaths in and just after the birth month is $472/417 = 11\%$ more than the number expected.

We can see that the death dip and death rise for sample 1 are larger than the death dips and death rises for the other three samples examined. In the month before the birth month, sample 1 has 45% fewer deaths than independence leads us to expect. In the remaining three samples combined, we observe 70 deaths in the month before the birth month; given independence, we expect about 75.25 $[(1/12) \times 903]$ in this period. Thus, just before the birth month, the observed number of deaths in samples 2, 3, and 4 combined is approximately 7% less than the number expected under independence.

Similarly, it is evident that the death rise in sample 1 is larger than the death rises in the remaining samples. We estimate that in sample 1, in the birth month and in the three months thereafter, there are 20% more deaths than expected. The equivalent figure for samples 2, 3, and 4 is 10%.

The death dip and death rise in the sample 1 may be larger than in the other samples because the members of sample 1 are considerably more famous than the members of the other samples. The 348 people in sample

1 are supposed to be the most famous people in American history. The larger number of people in the remaining samples were less stringently selected.

## RELATION BETWEEN FAME AND THE SIZE OF THE DEATH DIP AND DEATH RISE

We have referred several times to the notion that a group of famous people is expected to produce a larger death dip before the birth month than a group of ordinary people. Now we shall assess this idea more carefully. We classify the members of sample 1 into three groups according to how famous they are and examine the sizes of the death dip and death rise produced by each of these groups. If we are right, we should find that the more famous a group is, the larger its death dip and death rise.

There are obviously many ways to classify groups by fame. The method used here is convenient and seems plausible. The best-known members of "the four hundred" are those whose names have become household words in the "common culture." The "common culture" may be said to consist of the knowledge shared by almost all the members of a society—in other words, some sort of lowest common denominator of knowledge. To find which of the four hundred is "in" the common culture, we must find a set of people who know only what is in that culture. The members of the four hundred who are known to this group of people have names that are part of the common culture.

Of all the people in a society, children come closest to having no more knowledge than is in the common culture. If a child has heard of someone in the four hundred, he is very famous indeed. Thus the members of the four hundred who appear in children's biographies may be judged to be better known than members who do not appear in such biographies.

Two series of children's biographies were examined: Dodd, Mead's (1966) and Bobbs-Merrill's (1966). The criterion of coverage or noncoverage in these series can be used to classify members of the four hundred into groups of differing fame. Three different subgroups were formed from the original four hundred.

Group 1 consists of those of the four hundred whose names are found in both of the children's biography series. For example, George Washington, Thomas Jefferson, Benjamin Franklin, Mark Twain, and Thomas Edison are in group 1.

Group 2 consists of those of the four hundred whose names are in only one of the series. For example, John Quincy Adams, John Hancock, Jefferson Davis, Edgar Allen Poe, and Alexander Graham Bell are in group 2.



Group 3 consists of those of the four hundred whose names are in neither series. For example, Samuel Adams, Millard Fillmore, Rutherford B. Hayes, H. L. Mencken, and Nikola Tesla are in group 3.

For our purposes, the members of group 1 are judged to be more famous, on the average, than the members of group 2, and the members of group 2 are judged to be more famous, on the average, than the members of group 3. We say "on the average" because single individuals could be moved readily from one group to another if we used a third or fourth biography series to judge fame. But we think that the groups as a whole are ordered with respect to fame in the way we want them to be.

We can now measure the death dip and death rise produced by each of these groups. Table 5 gives the relevant information.

As predicted, the more famous a group is, the larger is its death dip. Group 1 produces a larger death dip than group 2, and group 2 produces a larger death dip than group 3. The death dip for the most famous group is quite large. About 78% of the deaths that are expected to fall in the month before the birth month do not do so. It should be stressed, however, that the number of people in group 1 is not large.

From Table 5 we can see also that the more famous the group is, the larger is the death rise that it produces. In group 1 (the most famous group) the observed number of deaths during and just after the birth month is about 58% greater than the number expected. Note that there is no death rise at all for group 3, the least famous group.

## SUMMARY

We have noted two sets of findings that are consistent with the notion that some people postpone death to witness a birthday because it is important to them. There is a death dip before the birth month and a death rise thereafter in four separate samples. We have noted also a consistent relation

TABLE 5. The Size of the Death Dip and Death Rise for Groups of Differing Fame

| GROUP | SIZE OF DEATH DIP (%) | SIZE OF DEATH RISE (%) | TOTAL NO. IN GROUP | NO. OF DEATHS IN THE BIRTH MONTH, AND 1, 2, 3 MONTHS THEREAFTER | NO. OF DEATHS IN THE MONTH BEFORE THE BIRTH MONTH |
|---|---|---|---|---|---|
| 1 | −78 | 58 | 55 | 29 | 1 |
| 2 | −63 | 23 | 129 | 53 | 4 |
| 3 | −20 | −3 | 164 | 53 | 11 |

between the fame of a group and the size of its death dip and death rise: the more famous the group, the larger the death dip and death rise it produces. These results might be due to chance, but this possibility is sufficiently small that we would prefer some other explanation of these phenomena.

There are indications that some people postpone dying in order to witness events other than their birthdays. There are fewer deaths than expected before the Jewish Day of Atonement in New York, a city with a large Jewish population. In addition, there is a dip in U.S. deaths, in general, before U.S. Presidential elections.

We have anecdotal evidence that the timing of death might be related to other important social events. For example, many people have noted that both Jefferson and Adams died on July 4th, 50 years after the Declaration of Independence was signed. We may find it easier to believe that this is not coincidental if we read Jefferson's last words, quoted by his physician.[1]

> About seven o'clock of the evening of that day, he [Jefferson] awoke, and seeing my staying at his bedside exclaimed, "Oh Doctor, are you still there?" in a voice however, that was husky and indistinct. He then asked, "Is it the Fourth?" to which I replied, "It soon will be." These were the last words I heard him utter.

## PROBLEMS

1. Why does the article examine the deaths of famous people only?

2. State and describe the two hypotheses being tested.

3. For the data in Table 1 is the assumption that "1/12 of the deaths occur each month" reasonable? Give reasons.

4. Calculate the expected number of deaths in each of the twelve starred cells in Table 1 under the assumption of independence, using the first method described in the text. Check that the expected total number of deaths in the month before the birth month is 28.3.

5. When would the two methods of estimating the total expected number of deaths in the month just before the birth month give exactly the same answer?

6. Refer to Figure 1.
   a. What is the meaning of the dotted line?
   b. Read from the graph the approximate number of deaths 1.5 months before the birth month. Is this figure meaningful? Why or why not?

[1] Merrill Peterson, *Thomas Jefferson and the New Nation* (New York: Oxford University Press, 1970), p. 1008.



   c. Suppose that the death month of famous people is really independent of the birth month. Would you then expect the graph to be a horizontal line? Why or why not?

7. Consider the following alternative explanation (which might be called the birthday bash theory) for the death rise during and after the birth month:

   "The reason for a death rise during these specific months is that famous people tend to exhaust themselves during the heavy celebrations on their birthdays thereby increasing the chance of death shortly after that date."

   From the data presented in this article can one distinguish between this explanation and the one proposed in the article?

8. Do you think sample 4 is an appropriate sample of the people whose names appear in the 1897–1942 volume of "Who Was Who" and whose surnames appear in "Foremost Families"? Comment.