Methods for Studying Coincidences

PERSI DIACONIS and FREDERICK MOSTELLER*

This article illustrates basic statistical techniques for studying coincidences. These include data-gathering methods (informal anecdotes, case studies, observational studies, and experiments) and methods of analysis (exploratory and confirmatory data analysis, special analytic techniques, and probabilistic modeling, both general and special purpose). We develop a version of the birthday problem general enough to include dependence, inhomogeneity, and almost and multiple matches. We review Fisher's techniques for giving partial credit for close matches. We develop a model for studying coincidences involving newly learned words. Once we set aside coincidences having apparent causes, four principles account for large numbers of remaining coincidences: hidden cause; psychology, including memory and perception; multiplicity of endpoints, including the counting of "close" or nearly alike events as if they were identical; and the law of truly large numbers, which says that when enormous numbers of events and people and their interactions cumulate over time, almost any outrageous event is bound to occur. These sources account for much of the force of synchronicity.

KEY WORDS: Birthday problems; Extrasensory perception; Jung; Kammerer; Multiple endpoints; Rare events; Synchronicity.

. . . for the 'one chance in a million' will undoubtedly occur, with no less and no more than its appropriate frequency, however surprised we may be that it should occur to *us*.

R. A. Fisher (1937, p. 16)

1. INTRODUCTION

Coincidences abound in everyday life. They delight, confound, and amaze us. They are disturbing and annoying. Coincidences can point to new discoveries. They can alter the course of our lives; where we work and at what, whom we live with, and other basic features of daily existence often seem to rest on coincidence.

Let us begin with a working definition. A coincidence is a surprising concurrence of events, perceived as meaningfully related, with no apparent causal connection. For example, if a few cases of a rare disease occur close together in time and location, a disaster may be brewing.

The definition aims at capturing the common language meaning of coincidence. The observer's psychology enters at *surprising*, *perceived*, *meaningful*, and *apparent*. A more liberal definition is possible: a coincidence is a rare event; but this includes too much to permit careful study.

Early in this century, biologist Paul Kammerer (see Sec. 2) and psychiatrist C. G. Jung separately studied coincidences and tried to develop a theory for them. Our definition is close to that given by Jung, who was deeply concerned with coincidences. Jung wrote a book called *Synchronicity: An Acausal Connecting Principle* (Jung 1973). Jung argued that meaningful coincidences occur far more frequently than chance allows. To account for this, he postulated synchronicity as "a hypothetical factor equal in rank to causality as a principle of explanation" (p. 19). Jung's images have captured the popular imagination, and synchronicity has become a standard synonym for coincidence.

We have organized this article around methods of studying coincidences, although a comprehensive treatment would require at least a large monograph.

2. OBSERVATIONAL STUDIES

One way to study coincidences is to look at the world around us, make lists, and try to find patterns and structure. Kammerer summarized years of work on such studies in his book *Das Gesetz der Serie: Eine Lehre von den Wiederholungen im Lebens—und im Weltgeschehen* (*The Law of Series: A Doctrine of Repetitions in Events in Life and Society*) (Kammerer 1919). Kammerer's traumatic career is brilliantly summarized in Arthur Koestler's (1971) book *The Case of the Midwife Toad*. This book includes a discussion and some translations of Kammerer's laws of series.

Kammerer is struck with the seeming repetition of similar events. He reports entries from his journal over a 15-year period. The following examples seem typical.

My brother-in-law E. V. W. attended a concert in Bosendorfer Hall in Vienna on 11 Nov. 1910; he had seat #9 and also coatcheck #9. (Kammerer 1919, p. 24)

On the walls of the Artist's Cafe across from the University of Vienna hang pictures of famous actors, singers, and musicians. On the 5th of May 1917, I noticed for the first time a portrait of Dr. Tyvolt. The waiter brought me the New Free Press, in which there was an article on the crisis in the German Popular Theater, with Dr. Tyvolt as the author. (Kammerer 1919, p. 27)

We would classify these examples as anecdotes. They seem a bit modest to be called case studies, but by making a collection Kammerer has moved them into the category of *observational* studies. An observational study consists of data collected to find out what things happen or how often events occur in the world as its stands, including comparisons of outcomes for different naturally occurring groups (Hoaglin, Light, McPeek, Mosteller, and Stoto 1982, pp. 6, 8, and 55–75).

Kammerer reported some 35 pages of examples. He

* Persi Diaconis is Professor of Mathematics, Department of Mathematics, and Frederick Mosteller is Roger I. Lee Professor of Mathematical Statistics Emeritus, Departments of Statistics, Biostatistics, and Health Policy and Management, Harvard University, Cambridge, MA 02138. This article was presented as the R. A. Fisher Memorial Lecture at the Joint Statistical Meetings in San Francisco, August 19, 1987. The authors express their appreciation to David Aldous, David Hoaglin, Ray Hyman, Augustine Kong, Bruce Levin, Marjorie Olson, Laurel Smith, and Cleo Youtz for their contributions to the manuscript. This work was facilitated in part by National Science Foundation Grants SES-8401422 to Harvard University and DMS86-00235 to Stanford University.

Journal of the American Statistical Association
December 1989, Vol. 84, No. 408, Applications & Case Studies

then gave a rough classification of series into types. His classifications were all illustrated by examples. To give but one, under "order within order" Kammerer wrote, "I arrange some objects in some way or other, play around with them, move them back and forth, and the largest come to rest together. Or, I sort some materials by their contents but the keywords arrange themselves in alphabetical order, without my doing anything" (Kammerer 1919, p. 51).

We have not succeeded in finding a clearly stated law of series in Kammerer's writing. The closest we have come is Koestler's translation "We thus arrive at the image of a world-mosaic or cosmic kaleidoscope, which, in spite of constant shufflings and rearrangements, also takes care of bringing like and like together" [translated from Kammerer (1919, p. 165) in Koestler (1971, p. 140)].

Modern Observational Studies in Psychology Related to Studying Coincidences. The brain processes and recalls information in ways that we barely understand. Clearly memory failure, selective attention, and the heuristic shortcuts we take in dealing with perceptions can sometimes deceive us into being surprised or lull us into ignoring rare events.

The literature offers some work on the psychology of coincidence. Some of this work relies on observational studies, and some of it verges into experiments when the investigator compares opinions on the same topic following different stimuli. In the interests of continuity we report such work in this section.

Over the past 10 years Ruma Falk and her collaborators have made a focused effort to study peoples' reactions to coincidences. In Falk (1982), she showed that the way a story is told can change its degree of surprise. Adding specific, superfluous details made the same coincidence seem more surprising. Surprise ratings increased if the same stories were told as potential future events as opposed to things that had just happened. The results suggest that upon hearing somebody's coincidence story, one is aware of a wide range of possibilities and considers the coincidence as one of many events that could have happened.

Falk (in press) and Falk and MacGregor (1983) showed that people found stories that happened to themselves far more surprising than the same stories occurring to others. Both of these findings agree with common sense; however, we all believe some things that are wrong. This substantial body of careful quantitative work is a fresh breeze in a sea of talk.

On a different theme, Hintzman, Asher, and Stern (1978) studied selective memory as a cause of coincidence. They claimed that we subconsciously keep track of many things. When a current event matches anything in memory, that event is retrieved from memory, and the coincidence remembered. We may thus overestimate the relative rate of coincidences.

Hintzman et al. (1978) set up a laboratory investigation involving free recall of incidentally learned words, which provides some evidence for the speculations. Falk (in press) reported that Kallai (1985) replicated their findings using events instead of words.

The probability problems discussed in Section 7 make the point that in many problems our intuitive grasp of the odds is far off. We are often surprised by things that turn out to be fairly likely occurrences. The body of work of Kahneman, Tversky, and their colleagues explores the way that people assign odds in the face of uncertainty. That work is presented and reviewed in Kahneman, Slovic, and Tversky (1982) and by Nisbett and Ross (1980).

The studies just described touch on a handful of important themes. They do not yet provide a broad, solid base for understanding the psychology of coincidences. Nevertheless, psychological factors are a main cause of perceived coincidence, and any systematic study, even if informal, is valuable.

Two related methods of studying coincidences are anecdotes and case studies. Unfortunately, space considerations force us to skip them.

3. EXPERIMENTS

Often investigators conduct experiments to test the powers of coincidence. We regard experiments as investigations where the investigator has control of the stimuli and other important determinants of the outcomes. We do not use *experiment* as in Harold Gulliksen's joke "They didn't have any controls because it was only an experiment."

Many experiments have been performed as tests of extrasensory perception (ESP). We illustrate with a single widely publicized experiment of Alister Hardy and Robert Harvie, reported in their book *The Challenge of Chance* [Hardy, Harvie, and Koestler 1973 (although Koestler is a coauthor of the book, his contributions do not deal with the experiment)].

Their experiment attempts to amplify any signal from an agent in an ESP study by using many agents simultaneously. A typical phase of their experiment uses 180 sending subjects and 20 receiving subjects, all gathered in a large hall.

A picture was shown or drawn for the sending subjects who concentrated on it, trying to send it with their minds to the receivers. The receivers made a drawing or wrote a statement that told what they received. At the end of a minute, monitors collected the drawings from the receivers.

The investigators isolated the receivers in cubicles with masked sides and fronts that face the drawing. We do not know how secure the cubicles were from outside sounds such as whispering. The experiment used feedback. After every trial the receivers left their cubicles and viewed the target figure. They then returned to the cubicle for the next transmission.

Scorers counted close matches (coincidences) between receivers' drawings and the pictures sent as successes. In informal reports of this and similar experiments, we often see examples of close matches, and Hardy and Harvie presented many examples. Our reaction to such successes is "Maybe there is something to this. How can we test it?"

Hardy and Harvie conducted a permutation test as follows: The total number of receiver responses over a seven-week period was 2,112. Out of these, 35 (or 1.66%) were judged direct hits. To obtain control data, the investigators compared targets with responses randomly shown in other trials. If the hits were attributable to chance, then about the same proportion of hits should occur. To ensure that the feedback would not corrupt the comparison, the investigators compared targets with responses chosen at random from trials held before the target trial.

The experiment had 35 hits and the control 59; however, the number of trials were as 2 to 5. One standard binomial significance test conditions on the total number of hits as $94 = 35 + 59$. It asks for the probability of 35 or more hits in 94 trials when the probability of success is $2/7 = .286$. Binomial tables give .0444. A two-sided level is about .09.

The experiment offers no strong evidence for ESP or a hidden synchronous force. Looking at some of the closely matched drawings, it is tempting to speculate, as did Hardy and Harvie, that perhaps just a small fraction of hits was really due to telepathy. Another speculative thought:

> It would certainly seem that at any rate the majority of the results in the original experiments, which at first sight might have suggested telepathy, can now be just as well explained by the coincidental coming together in time of similar ideas as has been so well demonstrated in the control experiments. (Hardy et al. 1973, p. 110)

4. EXPLORATORY DATA ANALYSIS

When data emerge in an unplanned way, we may still profitably analyze them, even using significance tests, though we regard them as having been subjected to exploratory rather than confirmatory data analysis. Usually the results are hypothesis-generating rather than firm conclusions, partly because of heavy affliction with problems of selection and/or multiplicity often unappreciated by even the most perceptive investigator. We provide one instructive illustration.

When the Hardy–Harvie experiment led to no convincing evidence of telepathy, the investigators explored other features of the investigation, as any investigator would and should. Sometimes in a single trial two or more receivers would generate much the same drawing or idea, even though their drawings did not seem to have a relation to the target image. In the large groups with 20 receivers (with sets of receivers in the true experiment compared with equivalent numbers of artificially created control groups who did not have the same images), they found the following numbers of coincidences for equal numbers of opportunities: *true experiment*—107 pairs, 27 triples, 7 quadruples, 1 quintuple; *control*—131 pairs, 17 triples, 2 quadruples. In small groups with 12 receivers they found *experiment*—20 pairs; *control*—23 pairs.

The authors regarded these multiples, which means several people getting the same answer, not as hits on the target, but as a sort of finding. They thought that these coincidences should perhaps be regarded as a form of telepathy. In their book, they put the issue as a question, especially because the controls are producing coincidences at about the same rate as the experimental subjects. The authors noted that triple matches and quadruple matches are more frequent in the experiment.

Recall that one feature in the experiment might tend to produce such multiple agreements. After each drawing, all receivers immediately looked at the drawing or slide that they had just tried to match. Possibly this exposure could encourage a next train of thought that might be similar across the several receivers. That concurrence would, however, be torn apart in the control samples. That is why the investigators in retrospect thought that showing the feedback to the receivers immediately after the trial may have been a mistake in design, because we have no way to match that effect in the control part of the study. Though conjectural, this experience illustrates both the hypothesis generation and the hazards of conclusions from exploratory data analysis. Before they began the study, the investigators did not know that they would want to make this comparison.

5. CONFIRMATORY DATA ANALYSIS

Did Shakespeare use alliteration when he wrote poetry? This question shocks our friends, many of whom respond with scornful looks and lines like "full fathom five thy father lies."

The psychologist B. F. Skinner (1939, 1941) analyzed this question by looking at Shakespeare's sonnets and counting how many times the same sound occurs in a line. He compared such coincidences with a binomial null hypothesis. For example, consider the sound *s* as it appears in lines of the sonnets. Table 1 shows the frequency distribution of the number of *s* sounds in lines of Shakespeare's sonnets. (The parentheses give counts Skinner obtained when the same word in a line is not counted more than once.)

After comparing the observed and the expected counts, Skinner concluded that, although Shakespeare may alliterate on purpose, the evidence is that binomial chance would handle these observations. He says, "So far as this aspect of poetry is concerned, Shakespeare might as well have drawn his words out of a hat" (Skinner 1939, p. 191).

In this problem the noise clearly overwhelms any signal. We find it refreshing to be reminded that things that "we all know must be true" can be very hard to prove.

6. ANALYTICAL TECHNIQUES—FISHER'S CONTRIBUTIONS

Special analytical statistical techniques may be developed for any field. For an R. A. Fisher lecture on coin-

Table 1. Skinner's Analysis of Alliteration in Shakespeare's Sonnets

	Number of s sounds in a line					Total count
	0	1	2	3	4	
Observed	702	501	161	29 (24)	7 (4)	1,400 1,392
Expected	685	523	162	26	2	1,398

Source: Skinner (1939, p. 189).

cidences it seems appropriate to display a special method developed by Fisher himself.

Fisher (1924, 1928, 1929) wrote three short papers on scoring coincidences. He suggested a way to give partial credit for partial matches and worked out appropriate versions for two different null distributions.

Fisher used matching of ordinary playing cards, as in ESP testing, for his main example. If someone guesses at a card, he or she can get it exactly right; get the value, suit, or color right; or merely match as a spot or picture card. The categories can be combined in various ways such as right in color and picture categories.

As an example of the need for Fisher's technique, J. Edgar Coover (1917), who carried out the first systematic studies of ESP, kept track of numbers of exact matches of playing cards and of matches in suits, or of colors. But he had no way of handling these various degrees of agreement simultaneously.

Fisher decomposed the possible matches into disjoint categories and suggested $-\log p$ as a reasonable score, where p is the chance of the observed match or a better one. The scores are shown in Table 2. The row labels relate to suit. Then 0 stands for no match, C for color, and S for suit. The column labels relate to values. Here 0 stands for no match, R for rank (picture or spot), and N for number (or value). Any pair of cards matches in some cell. For example, the jack of hearts and queen of diamonds match in color and rank, giving rise to the match CR.

The table entries show $-\log_{10} p$, where p is the chance of the observed match or better. Note that 0 means no match or better, so the score for cell 00 is $\log_{10} 1 = 0$.

This computation assumes that we draw both cards at random. Fisher's second paper (1928) begins by observing that in usual ESP tests, one card is chosen by a guessing subject, and the second is chosen from a randomly shuffled deck of cards. Subjects guess in a notoriously nonrandom manner. Fisher once discussed with one of us the troubles associated with this nonrandomness when a radio station investigated ESP by getting readers to send in guesses at the name of a card to be drawn. He said, "What if the first card drawn is the king of spades?"

Fisher suggests a new scoring system. His idea uses a *conditional score* based on the probability of the observed match, conditional on the subject's guess. Conditioning eliminates the need to explicitly model the distribution of the subjects' guesses. For his example, the conditioning event depends on the subject's guess only through its rank. Thus Fisher must provide two tables instead of one. The first table lists scores appropriate if the guesser names picture cards. The second table lists scores appropriate if the guesser names spot cards. Both tables are normalized to have the same mean and variance.

Table 2. Fisher's Logarithmic Scoring for Various Degrees of Card Matching

	Value		
Suit	−0	−R	−N
0−	0	.190	1.114
C−	.301	.491	1.415
S−	.602	.793	1.716

NOTE: Suit—0−, no match; C−, color; S−, suit. Value— −0, no match; −R, rank (picture or spot); −N, number.
Source: Fisher (1924, p. 183).

In his exposition, Fisher offers sage advice about practical aspects of evaluating coincidences. For example, he discusses whether people should be more moved when an individual has correctly guessed five cards in succession or on five separate occasions. He prefers the latter because it is hard to know the conditions under which such miracles occur. He wants to protect against fraud and mentions rabbits and conjurers.

Remark 1. Fisher's basic idea of using the value of $-\log p$ as a score is now standard statistical practice. He used it most effectively in his work on combining p values across independent experiments. Fisher derived the logarithmic scoring rule from the principle that the difference in scores given to two correctly guessed events should depend only on the ratio of their probabilities.

Remark 2. In card guessing, the score would be applied to successive runs through a 52-card deck. Then, the permutation distribution of the sum of scores is a basic object of interest. We observe that this depends on the guesser's choices. If the guesser always names the queen of hearts, the total score has *no* randomness. The probability theory for Fisher's conditional scoring system in this and other applications can be derived from Hoeffding's combinatorial central limit theorem. Label the values of an n-card deck as $1, 2, \ldots, n$. Define an $n \times n$ array of numbers $a(i, j)$ as follows: The ith row is determined by the guess on the ith trial by setting $a(i, j)$ to the value of the score if j turns up. For a given permutation π of the deck, the total score is $\sum_{i=1}^{n} a(i, \pi(i))$. Hoeffding (1951) showed that this quantity has an approximate normal distribution. Hoeffding's result puts some restrictions on the $a(i, j)$ that rule out cases where all rows are constant. Bolthausen (1984) gave nonlimiting approximations with error bounds.

Remark 3. We do not know why Fisher chose numbered cards versus face cards for the conditioning; a bridge player might have preferred honor cards verus nonhonors. But for ESP purposes, the sets *numbers* and *face cards* seem like more homogeneous groups to us. Fisher says this illustration may guide applications to more complex situations.

Remark 4. It is tempting to apply Fisher's idea to other problems. The idea of computing conditional scores is similar to the ideas of skill scoring used to evaluate weather forecasters. Diaconis (1978), Diaconis and Graham (1981), and Chung, Diaconis, Graham, and Mallows (1981) developed these ideas and gave references.

Remark 5. We think Fisher's scoring system is potentially broadly useful and suggest an application to the problem of birthdays and deathdays. Is there a hidden synchronous force causing a person to be born and die at nearly the same time of year? To investigate this, one looks at people's birth and death dates and scores them for closeness. One problem is the choice of cutoff: What counts as a match within a day or a month? Fisher's idea gets around choosing a cutoff by assigning a score of $-\log p$, with p the chance that two points chosen at random are at least as close as the observed pair. Here, distance is measured around a circle (mod 365). Because of homogeneity, the score is just

$$-\log_{10}\{(1 + 2 \times \text{observed distance})/365\}.$$

Hoaglin, Mosteller, and Tukey (1985, chap. 1) reviewed several empirical studies of the correlation between birthday and deathday.

7. MODELING

We find it convenient to divide our discussion of probability or statistical modeling into a part based on general-purpose models and a part based on special-purpose models. As Erich Lehmann remarked in the 1988 Fisher lecture, the separation betweeen these two kinds of models is very gray. Indeed, we think it is likely largely a matter of effort. Most of us would start out thinking of the birthday problem as a special model, but with petting and patting, it gradually becomes a general-purpose model, as we illustrate in Section 7.1.

7.1 General-Purpose Models: Birthday Problems

We find the utility of the birthday problem impressive as a tool for thinking about coincidences. This section develops four versions.

Problem 1: The Standard Birthday Problem. Suppose N balls are dropped at random into c categories, $N \le c$. The chance of no match (no more than one ball) in any of the categories is

$$\prod_{i=1}^{N-1}\left(1 - \frac{i}{c}\right). \tag{7.1}$$

This formula is easy to calculate, but hard to think about. If c is large and N is small compared to $c^{2/3}$, the following approximation is useful. The chance of no match is approximately

$$\exp(-N^2/2c). \tag{7.2}$$

This follows easily from Expression (7.1), using the approximation $\log_e(1 - i/c) \doteq -i/c$.

To have probability about p of at least one match, equate (7.2) to $1 - p$ and solve for N. This gives the approximations

$$N \doteq 1.2\sqrt{c} \text{ for a 50\% chance of a match} \tag{7.3}$$

and

$$N \doteq 2.5\sqrt{c} \text{ for a 95\% chance of a match.} \tag{7.4}$$

[The 2.5 in (7.4) is 2.448, but rounding up seems best.] Thus, if $c = 365$, $N = 22.9$ or 23 for a 50% chance and about 48 for a 95% chance. As far as we know, the earliest mention of the birthday problem was made by Von Mises (1939).

Problem 2: Many Types of Categories. The first variant offers a quantitative handle on a persistent cause of coincidences. Suppose that a group of people meet and get to know each other. Various types of coincidences can occur. These include same birthday; same job; attended same school (in same years); born or grew up in same country, state, or city; same first (last) name; spouses' (or parents') first names the same; and same hobby. What is the chance of a coincidence of some sort?

To start on this, consider the case where the different sets of categories are independent and the categories within a set are equally likely. If the numbers of categories in the sets are $c_1, c_2, \ldots, c_k$, we can compute the chance of no match in any of the categories and subtract from 1 as before. The following fairly accurate approximation is useful. If k different sets of categories are being considered, the number of people needed to have an even chance of a match in some set is about

$$N \doteq 1.2\sqrt{1\bigg/\left(\frac{1}{c_1} + \frac{1}{c_2} + \cdots + \frac{1}{c_k}\right)}.$$

The expression under the square root is the harmonic mean of the c_i divided by k. If all c_i equal c, the number of people needed becomes $1.2\,(c/k)^{1/2}$ so that multiple categories allow coincidences with fewer people as would be expected. For a 95% chance of at least one match, the multiplier 1.2 is increased to 2.5 as in Expression (7.4). The preceding approximation for N is derived by using Expression (7.2) and the independence of the categories.

As an illustration, consider three categories: $c_1 = 365$ birthdays; $c_2 = 1{,}000$ lottery tickets; $c_3 = 500$ same theater tickets on different nights. It takes 16 people to have an even chance of a match here.

Problem 3: Multiple Events. With many people in a group it becomes likely to have triple matches or more. What is the least number of people required to ensure that the probability exceeds $\frac{1}{2}$ that k or more of them have the same birthday? McKinney (1966) found, for $k = 3$, that 88 people are required. For $k = 4$, we require 187.

Levin (1981) treated this problem using multinomial theory. If N balls are dropped into 365 boxes, what is the chance that the maximum number of balls in a box equals or exceeds k? Using a clever twist of the Bayes theorem, Levin gave an algorithm that allows exact computation. He kindly carried this computation out for us, obtaining the results in Table 3. Thus in an audience of 1,000 people the probability exceeds $\frac{1}{2}$ that at least 9 people have the same birthday. We fit a curve to these numbers and find for modest k (say, smaller than 20) that $N \doteq 47\,(k - 1.5)^{3/2}$ gives a good fit.

In unpublished work we have shown that the following approximation is valid for fixed k and large c, with c the number of categories. The number of people required to have probability p of k or more in the same category is approximately given by solving for N in

$$Ne^{-N/ck}/(1 - N/c(k+1))^{1/k} = \left[c^{k-1}k!\,\log_e\left(\frac{1}{1-p}\right)\right]^{1/k}. \tag{7.5}$$

Table 3. Number N Required to Have Probability Greater Than 1/2 of k or More Matches With 365 Categories (Bruce Levin)

k	2	3	4	5	6	7	8	9	10	11	12	13
N	23	88	187	313	460	623	798	985	1,181	1,385	1,596	1,813

We obtained this result by using an approximation suggested to us by Augustine Kong.

Here is an example of its use. A friend reports that she, her husband, and their daughter were all born on the 16th. Take $c = 30$ (as days in a month), $k = 3$, and $p = \frac{1}{2}$. Formula (7.5) gives $N \doteq 18$. Thus, among birthdays of 18 people, a triple match in day of the month has about a 50–50 chance.

Problem 4: Almost Birthdays. We often find "near" coincidences surprising. Let us begin with the basic birthday problem. With 23 people it is even odds that at least 2 have the same birthday. How many people are needed to make it even odds that two have a birthday within a day? Abramson and Moser (1970) showed that as few as 14 people suffice. With seven people, it is about 50–50 that two have a birthday within a week of each other.

Changing the conditions for a coincidence slightly can change the numbers a lot. In day-to-day coincidences even without a perfect match, enough aspects often match to surprise us.

A neat approximation for the minimum number of people required to get a 50–50 chance that two have a match within k, when c categories are considered, follows from work of Sevast'yanov (1972). The answer is approximately

$$N = 1.2\sqrt{c/(2k + 1)}. \qquad (7.6)$$

When $c = 365$ and $k = 1$, this approximation gives about 13 people (actually 13.2) compared with the answer 14 from the Abramson–Moser calculation.

All variants discussed in this section can be modified to allow unequal probabilities and mild dependencies. Although we have assumed equal probabilities of falling in the various categories, Gail, Weiss, Mantel, and O'Brien gave several approaches to exact computation and an application to a problem of detecting cell culture contamination.

7.2 Special-Purpose Models

In this section, we illustrate special modeling for a few situations where we are commonly struck by coincidences.

The New Word. Here is a coincidence every person will have enjoyed: On hearing or reading a word for the first time, we hear or see it again in a few days, often more than once. Similarly, if a friend buys a fancy brand of automobile, we notice more and more of them on the road. We explain these coincidences, using new words as an example.

Suppose that you are 28 years old and that you graduated from secondary school at age 18. Let us consider only the time since graduation from secondary school, 10 years. You have just recognized a new word. Thus our first estimate of the rate at which you recognize this word is once in 10 years. After you have been tuned in to it, the word recurs within 10 days. Thus the rate has gone up by a factor of about 365.

What are some of the potential sources of change? An actual change in the rate of the appearance of this word could occur because of some change in your behavior—you read a book on a fresh topic and some technical expression arises, such as *bit* (a basic amount of information). Thus your behavior has changed the rate. This source of appearances of the word should not be regarded as coincidence because it has a fairly obvious cause.

In a less than blatant situation, you first come across the new word in your regular reading, say in a newspaper or magazine. Again the word appears in a few days in your reading or work. One possible explanation is that the world has changed. A word or expression formerly not in common use has become newsworthy. Thus an expression like *Watergate* can take on special meaning and be seen frequently, whereas formerly it had no real meaning except as a name for a building complex. We do not regard this as synchronicity.

A third causal explanation is heightened perception. Here, your regular reading turns up a word such as *formication* in an article on an otherwise routine topic, and you discover what it means. You see it again soon. Very likely this word has been going by your eyes and ears at a steady low rate, but you have not noticed it. You are now tuned in to it, and where blindness used to be sits an eagle eye. Thus heightened perception offers a third source of coincidence.

An important statistical phenomenon could also make some contribution to the higher-than-expected frequency. Because of our different reading habits, we readers are exposed to the same words at different observed rates, even when the long-run rates are the same. For high-frequency words like *the*, the difference is not only small, but irrelevant for this discussion. We are dealing with rare words. Imagine a class of rare words, all of which have about the same long-run frequency. Then, in the simplest situation, imagine being exposed to them at their regular rate. Some words will appear relatively early in your experience, some relatively late. More than half will appear before their expected time of appearance, probably more than 60% of them if we use the exponential model, so the appearance of new words is like a Poisson process.

On the other hand, some words will take more than twice the average time to appear, about $\frac{1}{7}$ of them ($1/e^2$) in the exponential model. They will look rarer than they actually are. Furthermore, their average time to *reappearance* is less than half that of their observed first appearance, and about 10% of those that took at least twice as long as they should have to occur will appear in less than 1/20 of the time they originally took to appear. The model we are using supposes an exponential waiting time to first occurrence of events. The phenomenon that accounts for part of this variable behavior of the words is of course the regression effect.

We now extend the model. Suppose that we are somewhat more complicated creatures, that we require k exposures to notice a word for the first time, and that k is itself a Poisson random variable with mean $\lambda + 1$. In fact, for definiteness suppose $k - 1$ is distributed according to a Poisson distribution with mean $\lambda = 4$. What we are saying is that we must have multiple exposures before we detect a new word, and that the multiple varies from one word to another for chancy reasons.

Then, the mean time until the word is noticed is $(\lambda + 1)T$, where T is the average time between actual occurrences of the word. The variance of the time is $(2\lambda + 1)T^2$. Suppose $T = 1$ year and $\lambda = 4$. Then, as an approximation, 5% of the words will take at least time $[\lambda + 1 + 1.65\ (2\lambda + 1)^{1/2}]T$ or about 10 years to be detected the first time. Assume further that, now that you are sensitized, you will detect the word the next time it appears. On the average it will be a year, but about 3% of these words that were so slow to be detected the first time will appear within a month by natural variation alone. So what took 10 years to happen once happens again within a month. No wonder we are astonished.

One of our graduate students learned the word formication on a Friday and read part of this manuscript the next Sunday, two days later, illustrating the effect and providing an anecdote.

Here, sensitizing the individual, the regression effect, and the recall of notable events and the nonrecall of humdrum events produce a situation where coincidences are noted with much higher than their expected frequency. This model can explain vast numbers of seeming coincidences.

The Law of Truly Large Numbers. Succinctly put, the law of truly large numbers states: With a large enough sample, any outrageous thing is likely to happen. The point is that truly rare events, say events that occur only once in a million [as the mathematician Littlewood (1953) required for an event to be surprising] are bound to be plentiful in a population of 250 million people. If a coincidence occurs to one person in a million each day, then we expect 250 occurrences a day and close to 100,000 such occurrences a year.

Going from a year to a lifetime and from the population of the United States to that of the world (5 billion at this writing), we can be absolutely sure that we will see incredibly remarkable events. When such events occur, they are often noted and recorded. If they happen to us or someone we know, it is hard to escape that spooky feeling.

A Double Lottery Winner. To illustrate the point, we review a front-page story in the *New York Times* on a "1 in 17 trillion" long shot, speaking of a woman who won the New Jersey lottery twice. The 1 in 17 trillion number is the correct answer to a not-very-relevant question. If you buy one ticket for exactly two New Jersey state lotteries, this is the chance both would be winners. (The woman actually purchased multiple tickets repeatedly.)

We have already explored one facet of this problem in discussing the birthday problem. The important question is What is the chance that some person, out of all of the millions and millions of people who buy lottery tickets in the United States, hits a lottery twice in a lifetime? We must remember that many people buy multiple tickets on each of many lotteries.

Stephen Samuels and George McCabe of the Department of Statistics at Purdue University arrived at some relevant calculations. They called the event "practically a sure thing," calculating that it is better than even odds to have a double winner in seven years someplace in the United States. It is better than 1 in 30 that there is a double winner in a four-month period—the time between winnings of the New Jersey woman.

8. TOWARD A RATIONAL THEORY OF COINCIDENCES

The preceding review provides some examples of how to think about and investigate coincidences. This final section lists our main findings. Although we do not yet have a satisfactory general theory of coincidences, a few principles cover a very large measure of what needs to be explained. Here we sketch four principles.

Hidden Cause. Much of scientific discovery depends on finding the cause of a perplexing coincidence. Changes in the world can create coincidences; likewise, changes in our own behavior such as a new pattern of reading or eating can create a pattern. Frequency of forecasting the same dire event improves the chances of simultaneity of forecast and outcome. Forgetting many failed predictions makes success seem more surprising.

At the same time, vast numbers of coincidences arise from hidden causes that are never discovered. At the moment, we have no measure of the size of this body of occurrences. Similarly, we have no general way to allow for misrepresentation, mistaken or deliberate, that may lead to many reports of coincidences that never occurred.

Psychology. What we perceive as coincidences and what we neglect as not notable depends on what we are sensitive to. Some research suggests that previous experience gives us hooks for identifying coincidences. Multiple events emphasize themselves, and without them we have no coincidence to recognize. The classical studies of remembering remind us that frequency, recency, intensity, familiarity, and relevance of experience strengthen recall and recognition. Thus classical psychology has much to teach us about coincidences because they depend so much on recall and recognition.

Multiple Endpoints and the Cost of "Close." In a world where close to identity counts, as it is often allowed to do in anecdotes, and "close" is allowed to get fairly far away, as when Caesar spoke of a military victory as avenging the death in battle, 50 years earlier, of the grandfather of his father-in-law, as if it were a personal revenge (Caesar 1982, p. 33), then the frequency of coincidences rises apace. Some formulas presented here emphasize the substantial effect that multiple endpoints can have.

The Law of Truly Large Numbers. Events rare per person occur with high frequency in the presence of large numbers of people; therefore, even larger numbers of interactions occur between groups of people and between people and objects. We believe that this principle has not yet been adequately exploited, so we look forward to its further contribution.

Concluding Remarks

In brief, we argue (perhaps along with Jung) that coincidences occur in the minds of observers. To some extent we are handicapped by lack of empirical work. We do not have a notion of how many coincidences occur per unit of time or of how this rate might change with training or heightened awareness. We have little information about how frequency of coincidences varies among individuals or groups of people, either from a subjective or from an objective point of view. Although Jung and we are heavily invested in coincidences as a subjective matter, we can imagine some objective definitions of coincidences and the possibility of empirical research to find out how frequently they occur. Such information might help us.

To get a better grip on coincidences that matter to people, it might be useful to employ a critical incidence study. The results might help us distinguish between those coincidences that genuinely move people and those that they regard as good fun though not affecting their lives. Such distinctions, if they are valid, would help focus further coincidence studies on matters people think are important.

In a culture like ours based heavily on determinism and causation, we tend to look for causes, and we ask What is the synchronous force creating all of these coincidences? We could equally well be looking for the stimuli that are driving so many people to look for the synchronous force. The coincidences are what drive us. And the world's activity and our labeling of events generates the coincidences.

The more we work in this area, the more we feel that Kammerer and Jung are right. We are swimming in an ocean of coincidences. Our explanation is that *nature* and we ourselves are creating these, sometimes causally, and also partly through perception and partly through objective accidental relationships. Often, of course, we cannot compute the probabilities, but when we can, such computations are informative. Where we have solid control and knowledge, the rates of occurrences seem about as expected, as Fisher said, but our inexperience with and lack of empirical information about the kinds of problems coincidences present do make for many surprises.

[*Received December 1988. Revised April 1989.*]

REFERENCES

Abramson, M., and Moser, W. O. J. (1970), "More Birthday Surprises," *American Mathematical Monthly*, 77, 856–858.

Bolthausen, E. (1984), "An Estimate of the Remainder in a Combinatorial Central Limit Theorem," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 66, 379–386.

Caesar [Julius] (1982), *The Conquest of Gaul* (trans. S. A. Handford), Hammondsworth, U.K.: Penguin (original circa 50 B.C.).

Chung, F. R. K., Diaconis, P., Graham, R. L., and Mallows, C. L. (1981), "On the Permanents of Complements of the Direct Sum of Identity Matrices," *Advances in Applied Mathematics*, 2, 121–137.

Coover, J. E. (1917), *Experiments in Psychical Research at Leland Stanford Junior University* (Psychical Research Monograph 1), Stanford University, Leland Stanford Junior University.

Diaconis, P. (1978), "Statistical Problems in ESP Research," *Science*, 201, 131–136. (Reprinted in 1985 in *A Skeptic's Handbook of Parapsychology*, ed. P. Kurtz, Buffalo, NY: Prometheus, pp. 569–584.)

Diaconis, P., and Graham, R. (1981), "The Analysis of Sequential Experiments With Feedback to Subjects," *The Annals of Statistics*, 9, 3–23.

Falk, R. (1982), "On Coincidences," *The Skeptical Inquirer*, 6, 18–31.

——— (in press), "The Judgment of Coincidences: Mine Versus Yours," *American Journal of Psychology*.

Falk, R., and MacGregor, D. (1983), "The Surprisingness of Coincidences," in *Analysing and Aiding Decision Processes*, eds. P. Humphreys, O. Svenson, and A. Vári, New York: North-Holland, pp. 489–502.

Fisher, R. A. (1924), "A Method of Scoring Coincidences in Tests With Playing Cards," *Proceedings of the Society for Psychical Research*, 34, 181–185. [Reprinted in 1971 in *Collected Papers of R. A. Fisher* (Vol. 1), ed. J. H. Bennett, Adelaide, Australia: University of Adelaide, pp. 557–561.]

——— (1928), "The Effect of Psychological Card Preferences," *Proceedings of the Society for Psychical Research*, 38, 269–271. [Reprinted in 1972 in *Collected Papers of R. A. Fisher* (Vol. 2), ed. J. H. Bennett, Adelaide, Australia: University of Adelaide, pp. 241–242.]

——— (1929), "II: The Statistical Method of Psychical Research," *Proceedings of the Society for Psychical Research*, 39, 189–192. [Reprinted in 1972 in *Collected Papers of R. A. Fisher* (Vol. 2), ed. J. H. Bennett, Adelaide, Australia: University of Adelaide, pp. 388–391.]

——— (1937), *The Design of Experiments* (2nd ed.), London: Oliver & Boyd.

Gail, M. H., Weiss, G. H., Mantel, N., and O'Brien, S. J. (1979), "A Solution to the Generalized Birthday Problem With Application to Allozyme Screening for Cell Culture Contamination," *Journal of Applied Probability*, 16, 242–251.

Hardy, A., Harvie, R., and Koestler, A. (1973), *The Challenge of Chance: Experiments and Speculations*, London: Hutchinson.

Hintzman, D. L., Asher, S. J., and Stern, L. D. (1978), "Incidental Retrieval and Memory for Coincidences," in *Practical Aspects of Memory*, eds. M. M. Gruneberg, P. E. Morris, and R. N. Sykes, New York: Academic Press, pp. 61–68.

Hoaglin, D. C., Light, R. J., McPeek, B., Mosteller, F., and Stoto, M. A. (1982), *Data for Decisions: Information Strategies for Policymakers*, Cambridge, MA: Abt Books.

Hoaglin, D. C., Mosteller, F., and Tukey, J. W. (eds.) (1985), *Exploring Data Tables, Trends, and Shapes*, New York: John Wiley.

Hoeffding, W. (1951), "A Combinatorial Central Limit Theorem," *The Annals of Mathematical Statistics*, 22, 558–566.

Jung, C. G. (1973), *Synchronicity: An Acausal Connecting Principle* (trans. R. F. C. Hull; Bollingen Ser.), Princeton, NJ: Princeton University Press.

Kahneman, D., Slovic, P., and Tversky, A. (1982), *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge, U.K.: Cambridge University Press.

Kallai, E. (1985), "Psychological Factors That Influence the Belief in Horoscopes," unpublished master's thesis, Hebrew University, Jerusalem, Dept. of Psychology.

Kammerer, P. (1919), *Das Gesetz der Serie: Eine Lehre von den Wiederholungen im Lebens—und im Weltgeschehen*, Stuttgart: Deutsche Verlags-Anstalt.

Koestler, A. (1971), *The Case of the Midwife Toad*, New York: Random House.

Levin, B. (1981), "A Representation for Multinomial Cumulative Distribution Functions," *The Annals of Statistics*, 9, 1123–1126.

Littlewood, J. E. (1953), *A Mathematician's Miscellany*, London: Methuen. (Reprinted with minor corrections in 1957.)

McKinney, E. H. (1966), "Generalized Birthday Problem," *American Mathematical Monthly*, 73, 385–387.

Nisbett, R., and Ross, L. (1980), *Human Inference: Strategies and Short-*

comings of Social Judgment, Englewood Cliffs, NJ: Prentice-Hall.

Sevast'yanov, B. A. (1972), "Poisson Limit Law for a Scheme of Sums of Dependent Random Variables" (trans. S. M. Rudolfer), *Theory of Probability and Its Applications*, 17, 695–699.

Skinner, B. F. (1939), "The Alliteration in Shakespeare's Sonnets: A Study in Literary Behavior," *The Psychological Record*, 3, No. 15, 186–192.

——— (1941), "A Quantitative Estimate of Certain Types of Sound-Patterning in Poetry," *The American Journal of Psychology*, 54, 64–79.

Von Mises, R. (1939), "Über Aufteilungs—und Besetzungs-Wahrscheinlichkeiten," *Revue de la Faculté des Sciences de l'Université d'Istanbul, N.S.*, 4, 145–163. [Reprinted in 1964 in *Selected Papers of Richard von Mises* (Vol. 2), eds. P. Frank, S. Goldstein, M. Kac, W. Prager, G. Szegö, and G. Birkhoff, Providence, RI: American Mathematical Society, pp. 313–334.]