THE BIRTHDAY PROBLEM WITH UNLIKE PROBABILITIES

BY

KUMAR JOAG-DEV AND FRANK PROSCHAN

FSU STATISTICS REPORT #M798
The Birthday Problem With Unlike Probabilities

KUMAR JOAG-DEV and FRANK PROSHAN

University of Illinois
Department of Statistics
Champaign, IL 61820

and

The Florida State University
Department of Statistics
Tallahassee, Florida 32306-3033

AFOSR Technical Report No. 88-233

Department of Statistics, Florida State University,
Tallahassee, FL 32306-3033

February 1989

1 Research partially supported by AFOSR Grant No. 88-0040.
2 Research supported by AFOSR Grant No. 88-0040.

Key Words: Birthday Problem, Majorization, Schur Functions.

AMS Subject Classification: 62N05
ABSTRACT

In this note, we show that the notions of majorization and Schur functions can be introduced in a very simple way in an elementary course on probability in discussing the well known Birthday Problem when the probability of a birthday may vary with the day of the year. The probability that no 2 or more students share a birthday is readily shown to be a Schur-concave function of \((p_1, p_2, \ldots, p_{365})\), where \(p_i\) is the probability that when a birth occurs, it occurs on day \(i, i = 1, 2, \ldots, 365\).
Birthday Problem With Unlike Probabilities

KUMAR JOAG-DEV AND FRANK PROSCHAN

University of Illinois and Florida State University

1. Introduction. The Birthday Problem as described in Feller(1968, p.33) has become an important example in a course on elementary probability. The more inquiring student might ask "What happens when different dates of the year have different probabilities of being a birthday?" We know that the probability that no two or more students have the same birthday (called coincidence) is smaller in this case than in the standard case of all days being equally likely to be a birthday. Actually, the teacher may take this opportunity to teach a more detailed result: as the probabilities differ more and more from 1/365, the required probability decreases. More precisely, the desired probability is a Schur-concave function of the 365 probabilities.

The use and application of Schur-concavity has increased steadily since the publication of the excellent comprehensive book by Marshall and Olkin (1979). Unfortunately, it is still considered a research topic and too advanced to be used in elementary courses. Actually the basic notions of majorization and Schur functions are quite elementary when properly introduced. In this note, we explain these notions and their application to the Birthday Problem with unlike probabilities; we believe this explanation can be used successfully in teaching elementary probability.

Suppose we start with the simpler model of just two periods: the first half of the year (January 1 through June 30) and the second half (July 1 through December 31). The probability of being born in the first half is $p$ and in the second half is $1 - p$. With just two students in the class, the probability that they are born in different half-years is simply

$$P_2 = 2p(1 - p).$$

Note that this function of $p$ is symmetric about $1/2$. Furthermore, as $p$ and $1 - p$ move further apart, this probability decreases.

Next consider the Birthday Problem, where there are 365 days. Again we are interested in the event that no two or more students have the same birthday, in a class of \( n \) students. We assume now that the probability of being born on day \( i \) is \( p_i, \ i = 1, 2, \ldots, 365 \), where \( \sum_{i=1}^{365} p_i = 1 \). For simplicity, consider the simple case where \( n = 3 \). The probability that all three birthdays are different is

\[
P_{365} = \sum_{i \neq j \neq k} p_i p_j p_k, \tag{1}
\]

where summation is over all possible distinct choices of ordered triplets out of 365 days.

We wish to study the effect on \( P_{365} \) of increasing the spread or variability among the \( p_1, p_2, \ldots, p_{365} \). Specifically, suppose we start with \( 0 < p_1 < p_2 < 1 \) and then spread the \( p_1 \) and \( p_2 \) further apart while keeping their sum unchanged; the 363 remaining \( p_i \) are kept fixed. How does \( P_{365} \) change?

Note that the terms in (1) that contain neither \( p_1 \) nor \( p_2 \) as factors are unaffected. The terms that contain \( p_1 \) but not \( p_2 \) are of the form

\[
p_1 \sum_{i \neq j \neq k} p_i p_j. \tag{2}
\]

Similarly, the terms that contain \( p_2 \) but not \( p_1 \) are of the form

\[
p_2 \sum_{i \neq j \neq k} p_i p_j. \tag{3}
\]

Adding (2) and (3) we get

\[
(p_1 + p_2) \sum_{i \neq j \neq k \ 1 \ or \ 2} p_i p_j. \tag{4}
\]

Since we are keeping \( p_1 + p_2 \) constant, (4) is unchanged. Thus the only change occurs in the remaining terms \( p_1 p_2 \sum_{i \neq 1, 2} p_i \). But we have seen above that \( p_1 p_2 \) decreases as \( p_1 \) and \( p_2 \) move apart while their sum \( p_1 + p_2 \) remains fixed. Also \( \sum_{i \neq j \neq k \ 1 \ or \ 2} p_i p_j \) is unaffected by changes in \( p_1 \) or \( p_2 \). Hence \( P \) must decrease as \( p_1, p_2 \) move apart while \( p_1 + p_2 \) remains fixed.

This proves that for the case of \( n = 3 \), the probability \( P_{365} \) of no coincidence in birthdays decreases as the spread between a pair of \( p \) values increases. The proof for general \( n \) is similar.
Note that using different pairs of $p$ values, we can repeatedly spread the distance between the two elements of a pair keeping their sum fixed and achieve lower and lower values of $P$. It follows that among the set of $(p_1, p_2, \ldots, p_{365})$, the highest probability of no coincidence is achieved by $(1/365, 1/365, \ldots, 1/365)$ (having 0-spread), while $(1, 0, 0, \ldots, 0)$ has minimum $P$, (actually 0), and has the maximum spread.

Finally, we define majorization, Schur-concave and Schur-convex functions. A vector $x = (x_1, x_2, \ldots, x_n)$ is said to majorize a vector $x' = (x'_1, x'_2, \ldots, x'_n)$ if $x'$ can be derived from $x$ by a finite sequence of averagings. An averaging of $x$ yields $(x_1, x_2, \ldots, x_i, \ldots, x_j, \ldots, x_n)$, where $x'_i + x'_j = x_i + x_j$, while $|x'_i - x'_j| \leq |x_i - x_j|$. A function $f(x)$ is Schur-convex (Schur-concave) if $x$ majorizes $x'$ implies $f(x) \geq (\leq) f(x')$.

Note that the definition of majorization given just above is not the usual definition. The standard definition almost always used is:

Let $x[1] \geq x[2] \geq \cdots \geq x[n]$ denote a decreasing rearrangement of $x = (x_1, x_2, \ldots, x_n)$. Then $x$ majorizes $y$ if

$$\sum_{i=1}^{k} x[i] \geq \sum_{i=1}^{k} y[i] \quad \text{for } k = 1, 2, \ldots, n - 1,$$

and

$$\sum_{i=1}^{n} x[i] = \sum_{i=1}^{n} y[i].$$

We write $x \geq_{m} y$.

Although the two definitions are equivalent, the one we give in terms of averaging two elements at a time is more understandable and more directly usable to obtain results.

References