CLASSROOM CAPSULES

EDITOR

Michael K. Kinyon
Indiana University South Bend
South Bend, IN 46634

Classroom Capsules consists primarily of short notes (1–3 pages) that convey new mathematical insights and effective teaching strategies for college mathematics instruction. Please submit manuscripts prepared according to the guidelines on the inside front cover to Professors Ricardo Alfaro and Steve Althoen, University of Michigan-Flint, Flint, MI 48502.

The Birthday Problem Revisited

Eric Malm (emalm@stanford.edu), Stanford University, Stanford, CA 94305, Gail Nord (nord@gonzaga.edu), Gonzaga University, Spokane, WA 99258, James Colin Hill (c_hill@mit.edu), Massachusetts Institute of Technology, Cambridge, MA 02139, and John Nord (john.nord@sgs.org), Saint George's School, Spokane, WA 99208.

A standard problem [**1, 2**] frequently discussed in probability courses is to compute the number of people, n, that are needed to have at least a 50-50 chance that two or more of them share a birthday. This problem usually assumes that birth dates are uniformly distributed and all years have 365 days. The probability that in a group of n individuals at least two share a birthday is

$$\begin{aligned} P(\text{at least 2 people have the same birthday}) &= 1 - P(\text{all } n \text{ birthdays are distinct}) \\ &= 1 - \frac{365}{365}\frac{364}{365}\frac{363}{365}\cdots\frac{365-n+1}{365} \\ &= 1 - \frac{365!}{365^n(365-n)!}, \end{aligned}$$

which must be greater than or equal to one-half. Solving this transcendental inequality for the least integral value of n yields $n = 23$. This means in a group of 23 people there is at least a 50-50 chance that at least two of them share a birthday, but that date is not specified *a priori*.

This note modifies the standard problem by specifying a birthday; for this birthday, we choose a particular day, say January 1st. Computing the probability that, in a group of n individuals, at least one of them has their birthday on January 1st, we have

$$\begin{aligned} P\text{ (at least 1 person has a birthday on January 1st)} &= 1 - P\text{ (no birthdays fall on January 1st)} \\ &= 1 - \frac{364}{365}\frac{364}{365}\frac{364}{365}\cdots\frac{364}{365} \\ &= 1 - \left(\frac{364}{365}\right)^n. \end{aligned}$$

Suppose we wish to know the least n for which this probability is greater than or equal to $1/2$. Unlike the above problem, this has a closed form solution for n:

$$n > \frac{\ln \frac{1}{2}}{\ln \frac{364}{365}}.$$

The least integral value for which this holds is $n = 253$.

This problem can now be generalized to determine the probability of having exactly k people with a specific birthday in a group of size n, which we denote $P(n, k)$. To construct such a configuration, we first choose k people from the group of n to have January 1st as their birthday. The other $n - k$ individuals can have any of the other 364

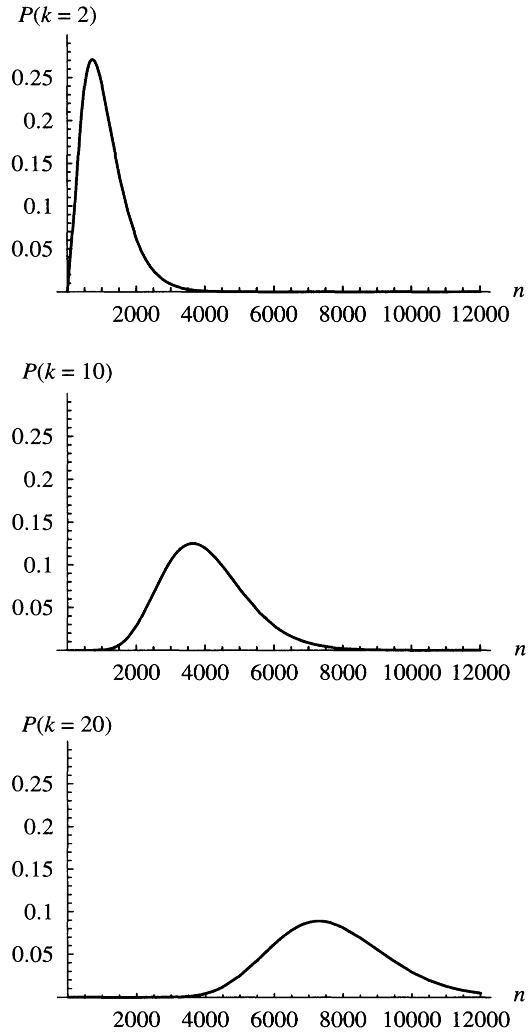

Figure 1. Plots of the probabilities of having exactly 2, 10, and 20 people with a January 1st birthday as a function of group size n.

days as their birthdays. Then

$$P(n,k) = \binom{n}{k}\left(\frac{1}{365}\right)^k \left(\frac{364}{365}\right)^{n-k},$$

which is easily seen to be a binomial distribution. See Figure 1 for plots of these probabilities for different values of k. These plots show the probabilities for specific values of k as a function of group size n. Figure 2 shows the probabilities as a function of group size n and the number of people k with January 1st birthdays. To generate these results for continuous n and k, we replace $n!$ in the formula for the binomial coefficients with its continuous interpolant $\Gamma(n+1)$ [**3**, p. 621].

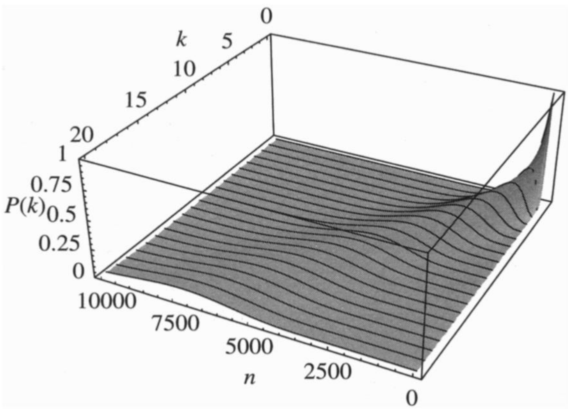

Figure 2. The probability that exactly k people in a group of size n have their birthdays on January 1st. The surface rulings represent integral values of k.

Figures 1 and 2 suggest that the maximum value of $P(n,k)$ for a fixed k occurs near $n = 365k$. We now demonstrate that this maximum value occurs at both $365k$ and $365k - 1$. Let $R(n,k)$ denote the ratio of $P(n+1,k)$ and $P(n,k)$, so that

$$R(n,k) = \frac{P(n+1,k)}{P(n,k)} = \frac{\binom{n+1}{k}\left(\frac{1}{365}\right)^k\left(\frac{364}{365}\right)^{n-k+1}}{\binom{n}{k}\left(\frac{1}{365}\right)^k\left(\frac{364}{365}\right)^{n-k}} = \frac{364(n+1)}{365(n-k+1)}.$$

For $n < 365k - 1$, we note that $364(n+1) > 365(n-k+1)$. Therefore, $R(n,k) > 1$, so $P(n+1,k) > P(n,k)$, and for these n-values the probabilities are increasing monotonically. Similarly, for $n > 365k - 1$, $R(n,k) < 1$, so the probabilities are decreasing monotonically. Finally, for $n = 365k - 1$,

$$R(365k-1,k) = \frac{364(365k)}{365(364k)} = 1,$$

so $P(365k-1,k) = P(365k,k)$, which are maxima by the monotonicity of the sequence above and below the corresponding n-values. We note that those seeking a simpler version of this generalization could restrict this analysis to the case where $k = 1$.

This capsule modifies the familiar birthday problem in ways that are accessible to students in a first course of probability or statistics. For further exploration, students can be directed to investigate birthday scenarios involving leap years, having birthdays fall within a range of dates, or choosing more than one birthday. Considering this last extension with k people having one birthday and ℓ having a different birthday provides an example with a multinomial distribution.

References

1. J. D. Austin, The Birthday Problem Revisited, *The Two-Year College Mathematics Journal*, 7 (1976) 39–42.
2. J. L. Devore, *Probability and Statistics for Engineering and the Sciences*, 4th ed., Duxbury Press, 1995.
3. W. Fulks, *Advanced Calculus: An Introduction To Anaylsis*, 3rd ed., Wiley, 1978.

A Geometric Look at Sequences that Converge to Euler's Constant

Duane W. DeTemple (detemple@wsu.edu) Washington State University, Pullman, WA 99164

In this article, we will generate and investigate sequences that converge to Euler's constant, γ. What is novel here is the elementary yet careful attention to the geometric descriptions of the terms of the sequences, allowing us to obtain a convergence rate of order $1/n^2$. Generally, this improved rate can be obtained only with a more cumbersome analytical analysis such as shown in **[2]**, **[3]**, and **[4]**.

Euler's constant is most often defined by comparing the natural logarithm, $\ln(n+1)$, with the nth partial sum

$$h_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$$

of the harmonic series $\sum_{k=1}^{\infty} 1/k$. As shown in Figure 1, h_n can be viewed as the sum of the areas of n rectangles of unit width and heights $1, 1/2, \ldots, 1/n$, and $\ln(n+1)$ is the area under the curve $y = 1/x$ over the interval $[1, n+1]$.

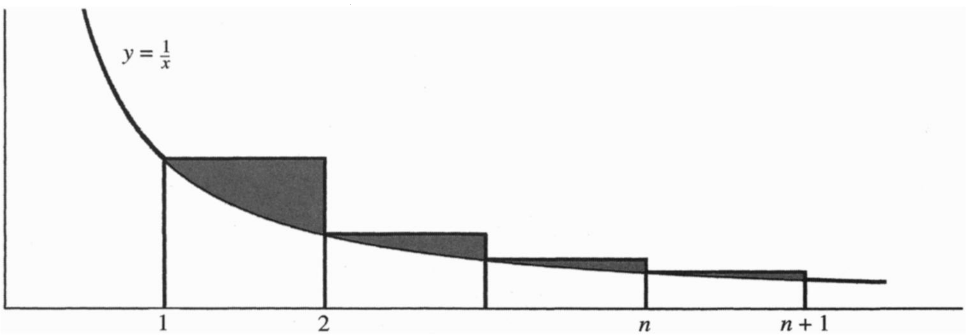

Figure 1.

The difference in these areas,

$$v_n = h_n - \ln(n+1) = 1 + \frac{1}{2} + \cdots + \frac{1}{n} - \ln(n+1),$$