BIRTHDAY GUESSING

PAUL C. KETTLER

ABSTRACT. This note investigates the estimation of probabilities that an individual has experienced a given birthday, based on the limited information of his mother's year of birth, and year of her life in which she gave birth. The distributions of each are assumed to be uniform and independent. This study was inspired by a guess of the age of a friend under these circumstances.

1. INTRODUCTION

Assume one knows the birth year of a mother, *e.g.*, 1961, and the year she gave birth to a son, *e.g.*, when she was 21, *viz.*, in her 22^{nd} year. Then, at some time in 2010, *e.g.*, one might naïvely assume that the son would be $2010-(1961+21)=28$ years of age. This well might be a reasonable approximation, but when, one could ask, would his 28^{th} birthday arrive within the year, or correspondingly, in looking at a particular date within 2010, what would the probability be that the son would have arrived at his 28^{th} birthday — or 27^{th} or 29^{th} birthday — the other two possibilities? This study provides an answer.

2. DEVELOPMENT

To generalize from the Introduction assume first a continuous time calendar, then let X be the fraction of a year within the mother's birth year b that she was born, and let Y be the fraction of a year following her j^{th} birthday that she gave birth to her son. Assume that X and Y are *i.i.d.* uniform, as defined, on the unit interval. Then $Z = X + Y$ is the fraction of a 2-year period beginning at the start of year $(b+j)$ that the son was born.

As is well known, Z has the triangular density $h(z) = 1 - |1 - z|$ on the interval $z \in [0, 2]$, where $h = f \star g$ is the convolution of the two respective uniform densities $f(x) \equiv 1$ and $g(y) \equiv 1$ with (x, y) on $[0, 1]^2$, or, in other words, if

$$H(z) = \iint\limits_{[0,1]^2} \mathbb{1}_{\{x+y \leqslant z\}} \, dx \, dy,$$

is the distribution of Z, then

$$h(z) = H'(z)$$

is its density.

Now, let a be the fraction in the 2-year period beginning with year n for which the question is asked, "How old is the son?". The probability $p_{n-(b+j+2)}$ that he has not yet reached

Date: 15 September 2010.
2010 *Mathematics Subject Classification.* Primary: 03B48. 44A35; Secondary: 47N30.
1991 *Journal of Economic Literature Subject Classification*. C16, C65.
Key words and phrases. Sum of independent random variables; Convolution.

his $n-(b+j+1)^{\text{th}}$ birthday is the mass in the right tail of $h(z)$ beginning with time $a+1$, in other words,

$$p_{n-(b+j+2)} = \int_{a+1}^{2} h(z)\,dz$$

Correspondingly, the probability $p_{n-(b+j)}$ that he has already reached his $n-(b+j)^{\text{th}}$ birthday is the mass in the left tail of $h(z)$ ending with time a, in other words,

$$p_{n-(b+j)} = \int_{0}^{a} h(z)\,dz$$

The residual, the probability $p_{n-(b+j+1)}$ that he is age $n-(b+j+1)$, is

$$p_{n-(b+j+1)} = 1 - \left(p_{n-(b+j+2)} + p_{n-(b+j)}\right)$$

For the given example, choosing $a = 0.6986$ in year $n = 2010$ (for the date 12 September 2010) the probabilities are as follows.

$$p_{26} = 0.0454$$
$$p_{27} = 0.7105$$
$$p_{28} = 0.2440$$

3. Conclusion

With diligence it is possible to refine the original naïve estimate, which was simply that the son was 28 years old on the given date. In this instance it was more probable that the son was 27, in fact his actual age.

(Paul C. Kettler)
CENTRE OF MATHEMATICS FOR APPLICATIONS
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OSLO
P.O. BOX 1053, BLINDERN
N–0316 OSLO
NORWAY
E-mail address: paulck@math.uio.no
URL: http://www.math.uio.no/~paulck/