The Actuary

April 1981 – Volume 15, No. 4

RATES AND PROBABILITIES ARE NOT EQUIVALENT TERMS

by Regina C. Elandt-Johnson

Ed. Note. Dr. Elandt-Johnson, of the Department of Biostatistics, School of Public Health, University of North Carolina, was a panelist at a Concurrent Session arranged by the American Statistical Association at our 1980 Annual Meeting in Montreal. Our invitation to her for this article resulted from a brief comment on this basic terminology question made at that session.

When I first became involved in analysis of mortality data, a small (or, perhaps, big?) problem which puzzled me was the use of the terms "rate" and "probability" by epidemiologists, demographers—and actuaries.

After reading several articles, and thinking a bit, I saw that there was some confusion in the use of these terms and thought I could easily demonstrate the differences and convince people to use each according to its appropriate meaning. So, I wrote an article published in the *Amer. J. Epidemiology 102* (1975), 3 [1], explaining the differences between these two concepts, as I saw (and see) them.

Although there were some positive responses, most people involved in epidemiological research (epidemiologists as well as statisticians) were not "converted." One of my colleagues, involved in teaching vital statistics, has pointed out that I might have been right to distinguish these terms, but people still use "rate" and "probability" in vital statistics as synonymous and it is not convenient to explain to the students why some old-established terms are now incorrect. I felt that this attitude was opportunism—hiding the head in the sand and letting things go on as they "always have been done," because it is too much trouble to put them right.

Actuaries, too, are still using these terms incorrectly, though from time to time they do wonder about their own definitions. Hence I welcome the opportunity to raise this question once more. I will open the discussion by presenting my point of view—as a mathematician and statistician—and inviting actuaries to respond.

Rates

If $y = y(x)$ is any mathematical function of x, then

$$\lim_{\Delta x \to 0} \frac{y(x + \Delta x) - y(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \frac{dy}{dx} = y'(x) \qquad (1)$$

is (conceptually) an instantaneous measure of change in y per unit change in x at the point x. It is called the *instantaneous absolute rate*. The quantity $\Delta y/\Delta x$ may be considered as an *average* rate over the short period Δx.

Thus, if the basic life table function l_x is represented by a continuous function of age x [$l_x = l(x)$], then the curve of death, $-dl_x/dx$, is formally the absolute rate function associated with the survival function l_x.

A more useful concept in describing chemical and biological processes, among others, is not the absolute, but the *relative* change per mass $\times$ time unit. If $y(x)$ represents a mathematical law according to which a certain mass decreases with time x, then the relative instantaneous rate per mass $\times$ time unit, at the time point x, is

$$\lim_{\Delta x \to 0} \left[-\frac{1}{y(x)} \frac{\Delta y}{\Delta x} \right] = -\frac{1}{y(x)} \frac{dy}{dx} = -\frac{d \log y(x)}{dx}. \qquad (2)$$

In this sense, the force of mortality

$$\mu_x = -\frac{1}{l_x} \frac{dl_x}{dx} = -\frac{d \log l_x}{dx} \qquad (3)$$

is a (relative) instantaneous death rate.

It is more difficult to obtain an average relative rate over the interval $(x, x + \Delta x)$, because we have to integrate the right-hand side of (2). In practice, however, we use the approximation

$$\int_x^{x+\Delta x} \frac{1}{y(t)} \frac{dy}{dt} \doteq \frac{1}{y(x')} \frac{\Delta y}{\Delta x}, \qquad (4)$$

where $x < x' < x + \Delta x$.

The corresponding average rate for a life table is the central rate, m_x, obtained from the formula

$$m_x = \frac{l_{x+1} - l_x}{L_x} = \frac{d_x}{L_x}. \qquad (5)$$

[For more details, see Elandt-Johnson and Johnson (1980) [2].]

Probability

The concept of probability is quite distinct from that of rate. It is concerned with stochastic phenomena and represents the chance of a certain event occurring. In particular, the event of interest may be death.

In terms of life table functions, the (cumulative) survival distribution function can be represented by l_x/l_0, and the familiar formula for the *conditional probability of death* between age x and age $x + 1$ given alive at age x is

$$q_x = \frac{l_{x+1} - l_x}{l_x} = \frac{d_x}{l_x}, \qquad (6)$$

or, alternatively,

$$q_x = 1 - \exp\left(- \int_0^1 \mu_{x+t} dt \right). \qquad (7)$$

Although formula (7) expresses a probability in terms of a rate, it does not mean that probability and rate are the same concepts.

For some unexplained (for me, anyway) reasons, in many actuarial books, q_x is called the "mortality rate" (as distinguished from the central death rate, m_x). Moreover, the conditional probability of surviving one year given alive at age x, $p_x = 1 - q_x$ is called the "survival" rate (!). How can one possibly speak about "survival rate"?

The confusion between rate and probability concepts arises because of time being involved. Clearly, for calculating prob-

(Continued on page 5)

Rates and Probabilities Are Not Equivalent Terms

(Continued from page 4)

abilities for living processes, it is necessary to introduce time, or equivalently age, x. But in this context, age is a stochastic variable with a certain distribution, so that the probability of dying over a period x to $x + 1$ involves that part of the probability distribution function over this interval. If x were to denote height, then the probability

$$\Pr\{\text{Height} > x + a \mid \text{Height} > x\}$$

would certainly not be called a rate!

Two Examples

To stimulate discussion, I quote from two actuarial texts in which some concepts of rates and probabilities are, in my opinion, a bit confusing.

(a) I first select the excellent book by Jordan (1967) [3] which I found of greatest value as a learner. Initially (in Chapter 1) he defines q_x as probability. However, in Chapter 14 on page 278, he says: "In the context of multiple-decrement table, $q_x^{\prime(k)}$ is solely a rate of decrement and must be distinguished from the probability $q_x^{\prime(k)}$. . . . In this book, the expression *rate of decrement* will always refer to the function $q_x^{\prime(k)}$ and will not be used as an abbreviation for *central rate of decrement* The function $q_x^{\prime(k)}$ has often been called the *absolute* rate in other actuarial literature."

But $q_x^{\prime(k)}$ is defined on page 277 by formula (14.30) and this has the same form as (1 13) on page 14, where it is defined as a probability. [In this article, formula (7).]

(b) In an interesting paper by A. H. Pollard (1980) [4], we read in §7.2 on page 243:

"The survival rate which includes deaths (!) (my mark of exclamation) from all causes is usually termed the observed survival rate. The risk of dying from causes of death other than the one under consideration varies with age. Comparisons of the survival experience of groups of patients that differ with respect to age and sex is made easier if the effect of mortality from other causes is eliminated. This is done by dividing the observed survival rate by the survival rate from deaths due to other causes. (Usually, for simplicity, mortality from all causes is used: this makes no significant difference.) The result is called the relative survival rate: it is the survival rate which would result if the cause of death under consideration were to be the only cause operating."

Although, I believe I am aware of reasons underlying the confusion in terminology, I do not think that I am able to understand this text entirely.

With only a small effort, terminology could easily be established describing concepts by appropriate names. I am looking forward to hearing some comments on this rather important matter.

References

1. Elandt-Johnson, R. C. (1975). Definition of Rates: Some Remarks on Their Use and Misuse. *Amer. J. Epidemiology 102*, 267–271.
2. Elandt-Johnson, R. C., and Johnson, N. L. (1980). *Survival Models and Data Analysis* J. Wiley and Sons, New York, Chapter 2.
3. Jordan, C. W. (1967). *Life Contingencies*. Society of Actuaries, Chicago, Chapters 1 and 14.
4. Pollard, A. H. (1980). The Interaction between Morbidity and Mortality. *JIA* 101, 233–302.

Letters

(Continued from page 3)

salary to each participant. Using, say, $100 a month and a 6% of final pay per year of service formula, this can easily be run on a standard valuation system for a plan providing $6 per year of service. As long as the increase in normal cost doesn't exceed the difference between 10- and 30-year amortization of the unfunded liability, the result can be used to determine contributions.

As Mr. Bader points out, even when these numbers aren't used to determine contributions, they give useful information to the plan sponsor. They can be produced economically using the above technique.

Matthew S. Easley

* * * *

Barnhart On Hunt

Sir:

Some response is in order to James H. Hunt's comments (January issue) about my cancer insurance article.

First, as to incomplete coverage: he calls cancer insurance, covering but one cause of loss, an absurdity. He may be surprised to learn that I agree with him. Personally, I wouldn't buy it.

But it doesn't follow that it is harmful or contrary to the public interest. Suppose my neighbor is bothered about, say, multiple sclerosis, and can buy insurance against it at a reasonable price, why should I try to deprive him of exercising that choice? Even more to the point, why should Mr. Hunt? Insurance regulatory attention grows less and less directed to protecting the public (and the industry) against harmful, unfair and unsound practices, and more and more toward mandating what, in the regulators' opinion, is best for the public—in some cases even to the point of mandating unsound practices and prohibitive costs!

But the key issue remains the loss ratio. I view loss ratios in terms of realistic present values of past and expected benefits vis-à-vis past and expected premiums, taking both interest and persistency into account. I can't follow the logic of Mr. Hunt's remark about guaranteed renewable policies, and I don't see that introducing non-forfeiture values would help; they would drive premium levels sharply higher and encourage still more lapsing.

Mr. Hunt misunderstood me in saying, "Mr. Barnhart can't be serious when

(Continued on page 6)