Gerontologia 8: 46–57 (1963)

Biometry Research Unit, Indian Statistical Institute, Calcutta

Age-Specific Probability Distributions of Annual Mortality*

By B. C. Das

Age has both biological and sociological implications. Biologically, it reflects the status of the organism, in terms of initial growth and subsequent decline; sociologically, it reflects the degree of the individual's participation in society as an earning, productive member and the degree to which other individuals may be dependent upon him. If all individuals born in the same year are examined according to age of death, it is seen that the likelihood of mortality is initially high in infancy, decreases through childhood, and then increases with advancing age. This general relationship represents a composite of several distinctly different relationships which are revealed when mortality is examined simultaneously in terms of age, sex, and cause of death. The value of these more precise expressions of age-specific mortality lies in the fact that they may point out causal factors otherwise not apparent, permit prediction of future events, and suggest lines for further research.

One approach to a more precise expression of age mortality relationships is the fitting of age-specific probability distributions to annual mortality differentiating between the sexes and causes of death. If the total deaths due to a particular cause are tabulated according to age of occurrence, then the number of deaths at each age may be converted to a proportion of the total death. Translation of the relationship between proportion of death and age into

* The work reported in this paper was supported in part by the U.S. Department of Health, Education and Welfare, Grant W-192 (C1), and in part by the National Institutes of Health, Grant RG 8569, and was carried out by the author at Case Institute of Technology, Cleveland, Ohio, USA.

Received for publication June 5, 1963.

a mathematical expression gives the age-specific probability distribution. Fitting this distribution to the observed data does not establish a causal relationship nor does it prove the existence of any natural law; it does describe the known facts quantitatively and does provide a mathematical model from which past and future events may be inferred (*Das*, 1962; *Das and Das*, 1962).

The age-specific probability distribution of annual mortality is dependent upon the age-sex structure of the reference population. The population structure of contemporary North American and Western European countries is similar. Four countries, whose population structures (*W.H.O.*, 1959) did not differ significantly have been taken for the present analysis: United States (white population only), United Kingdom (England and Wales), Sweden, and the Netherlands. In these four countries, the proportion of the population decreases from approximately .08 in the 0–4 age group to about .03 in the 70–74 age group; the decrease is linear with the exception of the proportions in the 20–24 and 25–29 age groups, which are relatively smaller than those of the adjacent age groups. The age-specific probability distributions of mortality for these reference populations characterize the actual amount of annual death by age and sex. Probability distributions have been fitted to data for five causes of death: motor vehicle accidents, cirrhosis of liver, vascular lesions affecting the nervous system, arteriosclerotic and degenerative heart disease, and malignant neoplasms. The observed data are reported in tables I to V for both sexes in the four countries analyzed. Three probability distributions have been fitted: gamma, normal and cumulative normal, and are illustrated for mortality among white males in the United States.

Age-specific probability of annual mortality due to motor vehicle accidents among white males in the United States (*W.H.O*, 1959) follows a gamma-distribution. The percentage distribution increases rapidly to age 20–24, after which it gradually declines. If the percentage of mortality, x, follows a gamma-distribution, its density function, f(x), is given by the following expression (*Mood*, 1950)

$$f(x) = \frac{1}{\alpha!\,\beta^{\alpha+1}} x^{\alpha} e^{-x/\beta} \qquad (1)$$

where α and β are the two parameters,

e is a constant, and
x is the variable quantity, age.

To illustrate the application of this distribution to motor vehicle accident deaths, a gamma-distribution has been fitted to the data in column (1) of table I and is given in figure 1. The fitted values for

Table I

Percentage Distributions of All Deaths by Age of Occurrence Due to Motor Vehicle Accidents in the USA, UK, Sweden and the Netherlands

Age	Male				Female			
	USA	UK	Sweden	Netherlands	USA	UK	Sweden	Netherlands
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
5–9	3.43	4.58	6.40	7.35	5.03	6.47	13.60	17.40
10–14	2.83	2.83	3.50	3.49	4.10	2.90	4.50	9.70
15–19	11.44	9.10	12.20	5.56	11.56	5.03	9.10	5.70
20–24	15.33	13.29	10.50	9.59	8.11	4.68	8.10	6.00
25–29	10.09	9.36	5.60	7.71	6.63	4.17	3.50	3.70
30–34	7.77	6.53	6.70	5.38	6.39	2.81	3.50	3.30
35–39	6.83	5.03	4.50	5.82	6.24	3.49	5.60	4.30
40–44	6.10	5.65	6.70	6.81	6.43	3.66	3.50	5.70
45–49	6.08	6.22	7.20	6.54	6.72	4.51	5.60	6.70
50–54	5.43	6.02	5.90	7.71	5.95	6.13	9.10	5.40
55–59	5.12	5.91	5.10	5.82	6.42	6.39	8.10	7.70
60–64	5.13	5.57	5.70	6.45	6.55	8.18	7.60	7.40
65–69	4.96	5.12	7.50	6.90	7.03	9.28	4.50	5.00
70–74	4.37	5.46	5.60	5.64	6.28	12.44	4.50	6.00
75–79	3.19	5.88	4.70	6.18	3.99	12.27	6.10	4.70
80–84	1.91	3.45	2.20	3.05	2.57	7.58	3.00	1.30

the parameters were $\alpha = .0949$ and $\beta = 2.8891$, permitting equation (1) to be written

$$f(x) = \frac{1}{(.0949)!\,(2.8891)^{1.0949}}\, x^{.0949}\, e^{-x/2.8891} \qquad (2)$$

This cause of death is influenced by the per capita use and number of motor vehicles, particularly automobiles. As the peak of mortality occurs during the early twenties, when the biological vigor of the individual is high, this probability distribution suggests that the causative factors are sociological and psychological.

Mortality due to chronic rheumatic heart disease and cirrhosis of the liver appears to be normally distributed according to age, that is, the percentage of mortality increases up to a certain age,

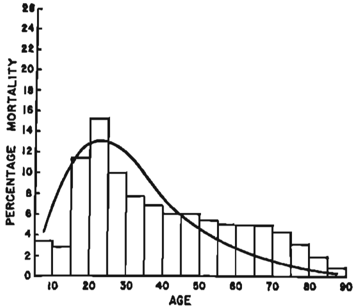

Fig. 1. Distribution of percentage of mortality by age due to motor vehicle accidents among white males in the United States. The histogram gives the observed data and the curve represents the fitted gamma-distribution. In this and subsequent figures, percentage mortality is indicated on the ordinate and age in years on the abscissa.

after which it decreases. If age-specific mortality, x, is normally distributed, its density function, f(x), is given by an expression of the form

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (3)$$

where π and e are constants,

μ and σ are the parameters, mean and standard deviation, and
x is the variable quantity, age (*Wadsworth and Bryan*, 1960).

To illustrate the interpretation of certain mortality processes as being normally distributed, normal curves have been fitted to the data for cirrhosis of liver for white males and females in the United States given in columns (1) and (5) of table II. In figure 2, the observed percentages of mortality are given by the histogram, and the fitted normal curve is superimposed upon the histogram. For males, the mean was 58.0710, indicating the age at which the maximum percentage of mortality occurs, and the standard deviation was 12.0635. Taking the mean, plus and minus one standard deviation, $\mu \pm 1\,\sigma$, or 58.0710 ± 12.0635, it can be stated that approximately 68% of the deaths would occur between the ages of 46.0075

Table II

Percentage Distributions of All Deaths by Age of Occurrence Due to Cirrhosis of Liver in the USA, UK, Sweden and the Netherlands

Age	Male				Female			
	USA	UK	Sweden	Netherlands	USA	UK	Sweden	Netherlands
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
5– 9	.04	.30	.80	1.00	.15	.20		.60
10–14	.05	.50	.40	1.00	.20	.40	.80	.60
15–19	.10	.30	.40		.26	.20	1.60	
20–24	.11	1.00			.30	.60	.80	.60
25–29	.36	.30		.50	.98	.20		
30–34	1.60	1.70	.40	1.00	2.59	.60		1.80
35–39	4.21	3.80	3.70	2.00	5.53	1.40	1.60	1.80
40–44	8.35	3.50	2.90	1.50	10.18	4.80	1.60	4.20
45–49	12.39	7.80	11.20	6.40	12.35	5.60	7.00	4.20
50–54	13.68	11.90	11.20	11.90	12.98	10.60	10.90	6.70
55–59	14.21	16.10	13.60	10.90	12.50	10.80	12.50	9.70
60–64	13.68	13.70	18.60	16.30	10.92	17.00	11.70	19.40
65–69	13.60	12.70	13.60	17.80	11.02	16.40	13.30	10.30
70–74	9.66	13.60	10.70	12.40	8.60	15.80	16.40	18.20
75–79	5.45	9.40	9.50	10.90	6.89	10.80	10.90	10.30
80–84	2.53	3.30	2.90	6.40	4.57	4.60	10.90	11.50

and 70.1345. For females, $\mu = 57.1275$, $\sigma = 13.2300$, and $\mu \pm 1\,\sigma$ corresponds to 43.8975 and 70.3575. Equation (3) could then be written

$$f(x) = \frac{1}{12.0635\sqrt{2\pi}} e^{-(x-58.0710)^2/2(12.0635)^2} \qquad (4)$$

for males, and

$$f(x) = \frac{1}{13.2300\sqrt{2\pi}} e^{-(x-57.1275)^2/2(13.2300)^2} \qquad (5)$$

for females. The expected age-specific mortality, as percentage of the total deaths falling at any age, x, would be given by computing f(x) from equation (4) or (5). To obtain the expected number of deaths, f(x) would be multiplied by the expected total number of deaths due to this cause, as computed from death rate data. Normally distributed mortality could possibly be interpreted as follows: while the incidence of the disease may be increasing in some manner with age, mortality attributed to the disease increases only up to a certain age (the mean), after which other diseases (those with a

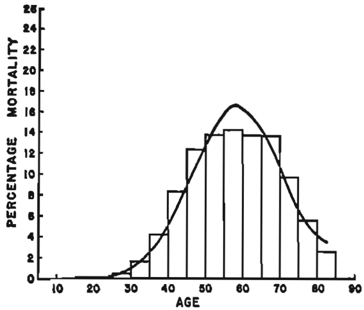

Fig. 2. Distribution of percentage of mortality by age due to cirrhosis of liver among white males in the United States. The histogram gives the observed data and the curve represents the fitted normal distribution.

cumulative mortality distribution) take precedence, with the result that mortality due to this disease appears to decrease. Diseases for which mortality shows a normal age-specific probability distribution take their toll in middle age, before the diseases of old age, but they may also cause a deteriorated condition predisposing individuals to collapse by diseases of old age.

Of considerable interest in countries where many of the contagious or epidemic causes of death have been controlled, and life expectancy is increasing, with a corresponding increase in the proportion of elderly individuals in the population, is the fact that mortality due to certain causes is an increasing function of age. Specifically, the percentage distribution of mortality is characterized by an ogive or cumulative normal distribution. The cumulative normal distribution is exhibited formally as (*Mood*, 1950)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-[(t-\mu)^2/2\sigma^2]}\, dt \qquad (6)$$

where f(x) is the cumulative probability density and t represents the variable quantity, age. For three major causes of death, cumulative normal distributions have been fitted to the data. The procedure involved converting the percentages of mortality to probits (normal deviates with $\mu = 5$ and $\sigma = 1$) (*Finney*, 1947), fitting a straight line to the probits by the method of least squares (*Lewis*, 1960), and converting the fitted probits back to percentages.

Designating expected mortality in probits as Y', the relationship with age, x, can be simply expressed by the linear equation

$$Y' = k + mx \qquad (7)$$

Equation (7) has been fitted to the data for the mortality of white males and females in the United States attributed to vascular lesions affecting the nervous system (table III), arteriosclerotic and degenerative heart disease (table IV), and malignant neoplasms (table V). Equation (7), with the appropriate values of m and k, is given below for the six sets of data.

Cause of death	Sex	Equation: $Y' = k + mx$	Equation number
Vascular lesions affecting	Male	$Y' = 1.2470 + .0382\,x$	(8)
nervous system	Female	$Y' = 1.2221 + .0378\,x$	(9)
Arteriosclerotic and degenerative	Male	$Y' = .7666 + .0478\,x$	(10)
heart disease	Female	$Y' = .4899 + .0489\,x$	(11)
Malignant neoplasms	Male	$Y' = 1.6998 + .0339\,x$	(12)
	Female	$Y' = 1.8795 + .0296\,x$	(13)

In order to compare equations (8) to (13), and their implications, it is appropriate to first note some properties of k and m. First, k is a constant and indicates the percentage (as a probit) of mortality at age 0. A high value of k indicates that mortality takes place even at the earliest ages; a low value of k indicates later onset of mortality. For these six equations, malignant neoplasms in females shows the highest initial mortality, while arteriosclerotic and degenerative heart disease are later in their onslaught. Second, m provides an indication of the rate at which mortality increases with age. A high m indicates a rapid increase in mortality with age, while lower values of m indicate a more gradual increase in mortality with age. The lowest value of m, for these six equations, is for malignant neoplasms in females; for both sexes, arteriosclerotic and degenerative heart disease show the highest rate of increase. It should also be noted that mortality for these data could be expressed in terms of equation (6), inserting appropriate values of μ and σ. For white males in the United States, figure 3 presents the observed mortality as a histogram with the fitted curve superimposed on the histogram for vascular lesions affecting the nervous system (table III). It would appear that diseases showing a cumulative probability distribution of mortality actually affect the organism in an accumulative manner,

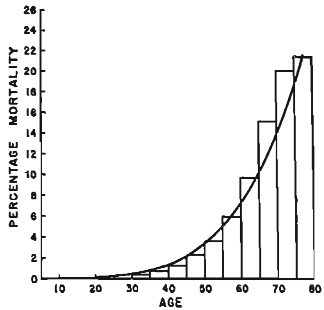

Fig. 3. Distribution of percentage of mortality by age due to vascular lesions affecting the nervous system among white males in the United States. The histogram gives the observed data and the curve represents the fitted cumulative normal distribution.

Table III

Percentage Distributions of All Deaths by Age of Occurrence Due to Vascular Lesions Affecting the Nervous System in the USA, UK, Sweden and the Netherlands

Age	Male				Female			
	USA	UK	Sweden	Netherlands	USA	UK	Sweden	Netherlands
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
5–9	.07	.03		.02	.07	.01		
10–14	.09	.07	.03	.05	.07	.05	.04	.07
15–19	.10	.06	.15	.11	.08	.04	.02	.04
20–24	.16	.13	.21	.10	.15	.07	.10	.02
25–29	.22	.17	.13	.30	.19	.11	.10	.24
30–34	.37	.26	.26	.27	.40	.21	.16	.28
35–39	.66	.44	.34	.34	.64	.40	.40	.32
40–44	1.24	.93	.80	.45	1.18	.77	.60	.52
45–49	2.24	1.88	1.73	.88	2.03	1.63	1.95	.96
50–54	3.60	3.40	3.20	2.20	3.04	3.00	2.47	2.02
55–59	5.96	6.14	4.66	3.47	4.80	4.77	4.95	3.80
60–64	9.72	9.67	8.74	7.37	7.83	7.78	7.62	6.71
65–69	15.15	14.84	12.99	12.13	13.15	12.69	14.22	12.22
70–74	20.05	19.93	19.82	20.74	18.92	20.10	20.48	20.80
75–79	21.44	23.83	25.34	26.80	23.64	25.83	25.10	27.88
80–84	18.95	18.24	21.62	25.02	23.83	22.59	21.86	24.11

Table IV

Percentage Distributions of All Deaths by Age of Occurrence Due to Arteriosclerotic and Degenerative Heart Disease in the USA, UK, Sweden and the Netherlands

Age	Male				Female			
	USA	UK	Sweden	Netherlands	USA	UK	Sweden	Netherlands
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
5– 9	.01				.01			
10–14	.01			.01	.01			
15–19	.02		.02	.01	.02		.02	.01
20–24	.04	.02		.02	.03	.01	.02	.04
25–29	.10	.07	.08	.09	.05	.02	.03	.04
30–34	.36	.21	.14	.25	.14	.10	.09	.07
35–39	1.05	.58	.30	.38	.31	.11	.23	.23
40–44	2.43	1.40	.88	1.23	.68	.31	.24	.39
45–49	4.59	3.10	2.20	2.70	1.41	.73	.70	.82
50–54	6.94	5.80	3.98	5.02	2.68	1.56	1.74	1.53
55–59	10.16	8.49	7.80	8.34	5.14	3.29	3.01	3.49
60–64	13.58	11.01	11.20	11.31	9.34	6.68	6.30	7.21
65–69	16.40	14.77	15.20	14.50	14.97	12.20	12.84	13.18
70–74	17.06	18.61	19.42	18.81	20.09	20.02	21.28	21.16
75–79	15.38	19.73	20.93	19.98	23.05	27.09	27.91	27.23
80–84	11.90	16.25	17.89	17.34	22.07	27.87	25.58	24.61

or, strike the individual when an accumulated amount of general deterioration has taken place. As *Jones* (1956, p. 281) points out, "Many known properties and functional capacities of the body are

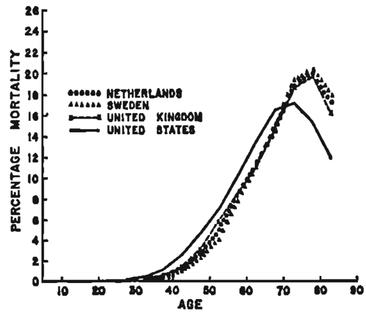

Fig. 4. Distributions of percentage of mortality by age due to arteriosclerotic and degenerative heart disease among males in the Netherlands, Sweden, United Kingdom, and United States.

Table V

Percentage Distributions of All Deaths by Age of Occurrence Due to Malignant Neoplasms in the USA, UK, Sweden and the Netherlands

Age	Male				Female			
	USA	UK	Sweden	Netherlands	USA	UK	Sweden	Netherlands
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
5–9	.62	.33	.46	.50	.58	.29	.29	.62
10–14	.38	.22	.24	.44	.33	.21	.22	.40
15–19	.40	.25	.38	.54	.31	.22	.26	.36
20–24	.45	.34	.22	.66	.36	.26	.31	.41
25–29	.68	.44	.69	.83	.71	.43	.58	.63
30–34	.98	.73	.86	1.14	1.56	1.08	1.30	1.36
35–39	1.48	1.21	1.22	1.48	2.71	1.80	1.97	2.34
40–44	2.67	2.49	2.09	2.54	4.57	3.64	3.82	3.96
45–49	4.56	4.87	3.17	4.16	6.81	6.06	5.48	6.14
50–54	7.37	8.48	6.66	7.07	8.64	8.25	7.95	7.83
55–59	10.94	12.60	8.96	10.41	10.80	10.57	9.91	10.30
60–64	14.99	14.76	12.66	13.89	13.12	12.76	12.90	12.55
65–69	16.99	16.51	16.28	14.83	14.57	14.45	15.06	14.00
70–74	16.02	16.02	17.38	16.32	14.10	15.58	15.73	15.04
75–79	13.01	13.30	16.92	14.85	12.15	14.60	14.41	13.72
80–84	8.46	7.46	11.81	10.34	8.68	9.82	9.80	10.32

changing – to the detriment of our existence – as accumulated time increases from birth.... It is likely... that disfunction accumulates and adds to the further tendency toward deterioration."

The percentage distributions of mortality due to arteriosclerotic

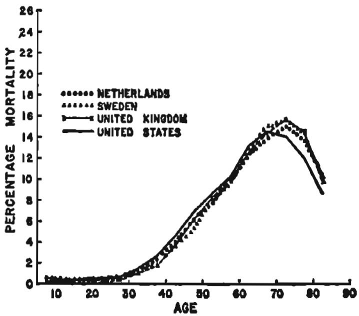

Fig. 5. Distributions of percentage of mortality by age due to malignant neoplasms among females in the Netherlands, Sweden, United Kingdom, and United States.

and degenerative heart disease for males and females in the United States, United Kingdom, Sweden and the Netherlands are given in table IV and are presented graphically in figure 4 for males. The distribution for United States males rises sooner and reaches a peak earlier than the distributions for the other three countries. Table V presents the corresponding data for malignant neoplasms, and figure 5 shows that the United States females die slightly earlier than their European counterparts due to this diagnosis. The mathematical expression of the age-specific distribution of mortality would be the same for all four countries, but the fitted parameters would differ. This finding implies that the underlying biological process, which the probability distribution may be reflecting, can be modified by environmental or sociological factors to yield varying parameter values.

Summary

Probability distributions have been fitted to the age-specific percentage of annual mortality according to cause of death. A gamma-distribution has been found to characterize death due to motor vehicle accidents among white males in the United States, with modal death occurring in the 20- to 24-year age group. A normal distribution describes mortality due to cirrhosis of liver in white males and females in the United States, with modal death occurring in the 55- to 59-year age group. A cumulative normal distribution is useful for describing mortality due to vascular lesions affecting the nervous system, arteriosclerotic and degenerative heart disease, and malignant neoplasms, in which modal death occurs after 64 years of age in both sexes. Percentage of mortality by age attributed to these five causes has been reported for both sexes in the Netherlands, Sweden, United Kingdom (England and Wales), and United States.

Zusammenfassung

Die wahrscheinliche Verteilung der altersspezifischen Prozente der jährlichen Sterblichkeit wurde mit der Todesursache in Übereinstimmung gebracht. Es wurde eine Gammaverteilung gefunden, welche die Todesfälle, die bei weißen Männern in den USA durch Verkehrsunfälle zustande kommen, charakterisiert mit zufälligem Tod im Alter von 20–24 Jahren. Eine normale Verteilung beschreibt die Sterblichkeit infolge Leberzirrhose bei weißen Männern und Frauen in den USA mit dem zufälligen Tod im Alter von 55–59 Jahren. Eine sich anhäufende normale Verteilung kann die Sterblichkeit infolge Gefäßstörungen, die das ZNS schädigen, arterioskleroticher und degenerativer Herzleiden oder infolge bösartiger Tumoren beschreiben, bei welchen ein zufälliger Tod im

Alter von über 64 Jahren bei beiden Geschlechtern vorkommt. Der Prozentsatz der Alterssterblichkeit mit diesen fünf Ursachen zusammen wurde für beide Geschlechter in Holland, Schweden, im United Kingdom (England und Wales) und in den USA genannt.

Résumé

Les distributions de probabilité ont été ajustées au pourcentage de la mortalité annuelle par classe d'âge, suivant les causes de décès. Les morts dûs à des accidents de circulation sont caractérisées par une distribution gamma, chez les sujets de race blanche de sexe masculin vivant aux U.S.A., la majorité des décès se situant entre 20 et 24 ans. La mortalité par cirrhose du foie chez les Blancs des deux sexes aux U.S.A. a une distribution normale, la majorité des décès survenant entre 55 et 59 ans. Une distribution cumulative est utile pour décrire la mortalité dûe aux lésions vasculaires touchant le système nerveux, ainsi qu'à l'artériosclérose, aux maladies dégénératives du système vasculaire et aux tumeurs malignes, pour lesquelles la majorité des décès survient après 64 ans dans les 2 sexes.

Le pourcentage de mortalité par âge pour ces 5 causes de décès est calculé pour les 2 sexes pour les Pays-Bas, la Suède, le Royaume Unis et les Etats Unis.

References

Das, B. C.: Age-specific changes in mortality, disability and demand on chronic disease hospitals. Systems Research Center Report SRC 21-M-62-5 (Case Institute of Technology, Cleveland, Ohio 1962).

Das, B. C. and *Das, Rhea S.:* Age-specific mortality and morbidity in relation to community health performance and demand on hospital facilities. Operations Research Group Report (Case Institute of Technology, Cleveland, Ohio 1962).

Finney, D. J.: Probit analysis (University Press, Cambridge 1947).

Jones, H. B.: A special consideration of the aging process, disease, and life expectancy; pp. 281–337, Vol. IV, Advances in biological and medical physics (Academic Press, New York 1956).

Lewis, D.: Quantitative methods in psychology (McGraw-Hill, New York 1960).

Mood, A. M.: Introduction to the theory of statistics (McGraw-Hill, New York 1950).

Wadsworth, G. P. and *Bryan, J. G.:* Introduction to probability and random variables (McGraw-Hill, New York 1960).

World Health Organization: Annual epidemiological and vital statistics 1956 (Geneva 1959).

Author's address: Dr. B. C. Das, Biometry Research Unit, Indian Statistical Institute, 203 Barrackpore Trunk Road, *Calcutta 35* (India)