Note on the Construction of Mortality Tables by Means of Compound Frequency Curves.*

BY

ARNE FISHER.

In the following pages I shall attempt to give a brief outline of a method of constructing a mortality table from the records of death by age and cause, but without knowing the exposed to risk at various ages. The method may to some appear new and revolutionary, and perhaps a few may assert it is impossible to construct a life table from such data alone without knowing the age distribution of the exposed to risk. It has, however, been done before, although by different methods than those I propose to employ.

The Danish Actuary, Dr. Phil. Jens Pedersen, in a brochure on "The Insurance of Substandard Lives" constructed special mortality tables for three separate danger classes of substandard risks, using certain Danish census data, giving the number of deaths in Danish cities and towns by cause and by age. Pedersen starts with certain hypotheses about the continuity of the danger regions and makes use of the differential geometry in his solution.

Karl Pearson, the eminent English biometrician, in a brilliant essay in his "Chances of Death" attacked the problem by frequency curve methods. Pearson took the numbers dying at various ages and analyzed the series into the sum of five frequency curves typical of old age, middle life, youth, childhood and infancy. Mr. Robert Henderson in his treatise on "Mortality and Statistics" claims that "it is difficult to lay a firm foundation for the method, because no analysis of the deaths into natural divisions by causes or otherwise has yet been made such that the totals in the various groups would conform to these frequency curves."

This statement evidently contains a half truth if we limit ourselves to the Pearsonian system of 7 types of frequency curves, which by no means suffice for all frequency distributions, occurring in practise. I have in fact in my statistical work encountered

* In part from a chapter of the author's forthcoming second volume of The Mathematical Theory of Probabilities (Macmillan Co., New York).

many statistical series which offer an absolute defiance to the Pearsonian curves. We will in general on the other hand succeed in fitting the curves if we instead of the Pearsonian method make use of the system of frequency curves originally introduced by Laplace and Poisson and of late years further developed by Charlier. I hope indeed in the following to show that it is possible to analyze the series of deaths into a system of 8 Laplacean-Charlier and Poisson-Charlier frequency curves, typical of distinctive groups of causes of deaths at various stages of life.

The Laplacean theory, the oldest and most general theory of frequency curves, assumes that any frequency curve is generated as the sum of different and independent frequency curves, generally infinite in number. It is on the basis of this hypothesis that I propose to analyze the series of deaths at various ages into 8 separate component frequency curves, typical of distinctive causes of death.

To start with fundamental principles let us consider the frequency of deaths at various ages out of an original cohort of say 10,000 persons, all entering under our observation at age 10. The question before us is: "How will those entrants gradually die off?" We will have a few deaths at age 10, these deaths slowly increase to age 20, then the increase becomes more rapid—although by no means uniform—until we reach the ages around 70, where we encounter a maximum. From age 70 and on the numbers of deaths rapidly decline, until at the age of 100 only a few survivors of the original cohort struggle along. Ultimately all are carried off by death at age 110 or so. Graphically this series—simply the d_x column of a mortality table—represents a compound curve with a high crest around age 70, gradually diminishing towards youth and old age. From 70 and on the curve resembles to a very high degree the right half of the Normal Laplacean or Gaussian Law of Error. This fact was already noticed by Lexis, who through an application of his dispersion theory analyzed the extreme right part as a normal curve with a maximum and mean in the neighborhood of age 70. Pearson went further and made a complete analysis of the remaining part of the curve.

My hypothesis is now that the numbers of deaths from specific causes cluster around certain definite ages in such a manner that the frequency distribution according to age from a specified cause or group of causes of death may be represented by a typical skew frequency curve of either type A (Laplace-Charlier Type) or type

B (Poisson-Charlier Type).* We will, for instance, notice that deaths from cancer and heart disease, etc., amongst the survivors at various ages of the original cohort are most frequent in the age period 65–70. Tuberculosis, on the other hand, levies its heaviest toll around the age interval 35–40. The first step is now to group the deaths arising from certain diseases in such a way that their sum total represents a frequency curve of either type A or type B. This classification or grouping is done by the aid of the Lexian-Charlier dispersion theory and by methods exactly like the ones described in my recent paper in the *Proceedings* (Vol. III, p. 241) on the construction of basic pure premiums. Through gradual tests I succeeded in getting the following 8 fairly typical groups.

B. Deaths typical of youth.
C. Deaths from industrial accidents, typhoid fever, etc.
D. Deaths from pulmonary tuberculosis, etc.
E. Deaths typical of middle life.
F. Deaths typical of late middle life.
G. Deaths typical of early old age.
H. Deaths typical of middle old age.
I. Deaths typical of extreme old age.

I do not claim those groups as being the final word. Probably the scheme might be further improved by decomposing group G

* The A type is given by the equation:

$$F(x) = \varphi(x) + \beta_3\varphi^{\mathrm{III}}(x) + \beta_4\varphi^{\mathrm{IV}}(x) + \cdots$$

where

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-(x-M)^2/2\sigma^2}$$

is the ordinary probability function expressed in units of the dispersion, ϕ^{III}, ϕ^{IV}, $\ldots$, its various derivatives, and M, σ, β_3, β_4 certain parameters, independent of the variate x; e is the logarithmic base. M is known as the mean, σ is commonly called the dispersion, β_3 the skewness and β_4 the excess.

The B type progresses by a finite difference formula of the following form: $F(x) = \psi_\gamma(x) + \gamma_2\Delta^2\psi_\gamma(x) + \gamma_3\Delta^3\psi_\gamma(x) + \cdots$, where the generating function now is:

$$\psi_\lambda(x) = \frac{e^{-\lambda}\lambda^x}{\underline{|x}},$$

or the Poisson exponential. λ and γ_2 are again certain parameters determined from the observations. The various parameters are throughout this paper computed by means of the well-known method of moments, that of least squares, or by the methods of Dr. Thiele's half invariants.

into at least two or three subgroups. A detailed and complete dispersion test would take considerable time, which at present, on account of other and more urgent work, is not at my disposal. However, the above grouping represents the most important factors in the generation of the total compound frequency curve, or the d_x column of the mortality table. I shall therefore proceed to apply the above classification to the deaths from the United States Registration Area for the year 1910 as recorded by age and cause of death in the volume "Mortality Statistics," published by the Federal Census Bureau, and construct a mortality table for white males for this area.

The computation of the various statistical parameters of the 8 curves by means of the well known method of moments gave the following results for the above mentioned material.

		Mean.	Dispersion.	Skewness.	Excess.
B type	B...........	Constructed by Jörgensen's method for quinquennial age			
	C...........	groups and interpolated for single ages by Novalis' formula.			
A type	D...........	44.3	16.306	−0.04465	−0.03344
	E...........	52.7	18.776	+0.02304	−0.02900
	F...........	59.7	17.110	+0.10812	−0.00705
	G...........	66.8	14.110	+0.15857	+0.05172
	H...........	71.3	14.587	+0.15992	+0.03773
	I...........	78.3	10.483	+0.15319	+0.06847

These parameters determine the equations for the various frequency curves $F(x)$, but it must be remembered that they give us only the form of distribution or the clustering tendency around the mean and no clue whatsoever as to the number of individual deaths in a particular group. In other words, we do not know the area of the curves themselves.

Let N_B, N_C, N_D, $\cdots$ N_I denote the areas of the various frequency curves and $N_A = 10{,}000$ be the area of the total compound frequency curve, i. e., the total number of deaths at various ages of quennial ages.

The total number dying at a particular age, say age x, will then according to our clustering hypothesis be:

$$d_x = N_B F_B(x) + N_C F_C(x) + N_D F_D(x) + \cdots + N_I F_I(x)$$

where

$$N_B + N_C + N_D + N_E + \cdots + N_I = N_A = 10{,}000.$$

In this equation we know the various values of $F(x)$ from the equation of the frequency curves, but none of the values of N, except $N_A = 10{,}000$.

Let us now for the present denote the proportionate mortality at a particular age or group of ages of any one of the above classes, say for instance class B, by $R_B(x)$. We have then for age x

$$R_B(x) = \frac{N_B F_B(x)}{N_B F_B(x) + N_C F_C(x) + \cdots N_I F_I(x)}.^*$$

Similar proportionate rates can be found for the other classes of causes of death. Moreover, it is readily seen that the various values of $R(x)$ for the different age groups depend upon the deaths alone in those groups and are independent of the number actually exposed to risk, provided, of course, that our numbers of deaths are so large that fluctuations due to sampling may be ignored. Bearing in mind that $N_B + N_C + N_D + \cdots + N_I = N_A = 10{,}000$, we have thus a larger number of observation equations than we have unknown and from which the values of the N's may be determined by well known methods. The best way is probably to determine first approximate values for the unknowns, and then adjust those values by the method of least squares so as to make as close a fit as possible with the observed values of $R(x)$.

The compound frequency curve, or what is the same, the d_x column of the mortality table, is simply found by addition of the above eight component curves. The result is shown graphically in Fig. 1. The areas of the 8 component skew frequency curves bear the following ratios to the total area:

B	.028140	F	.149095
C	.049857	G	.435717
D	.092627	H	.071631
E	.096724	I	.076209

The values of q_x are easily computed from the d_x column or the compound frequency curve by forming first a column of l_x by summing the series of death at various ages as represented by the compound curve. The well known formula $q_x = d_x / l_x$ gives us then the values of q_x.

* It is readily seen that $R_B(x)$ represents nothing more than the probability that when a person is dead the cause of death arose from the group of causes which we have designated by B.

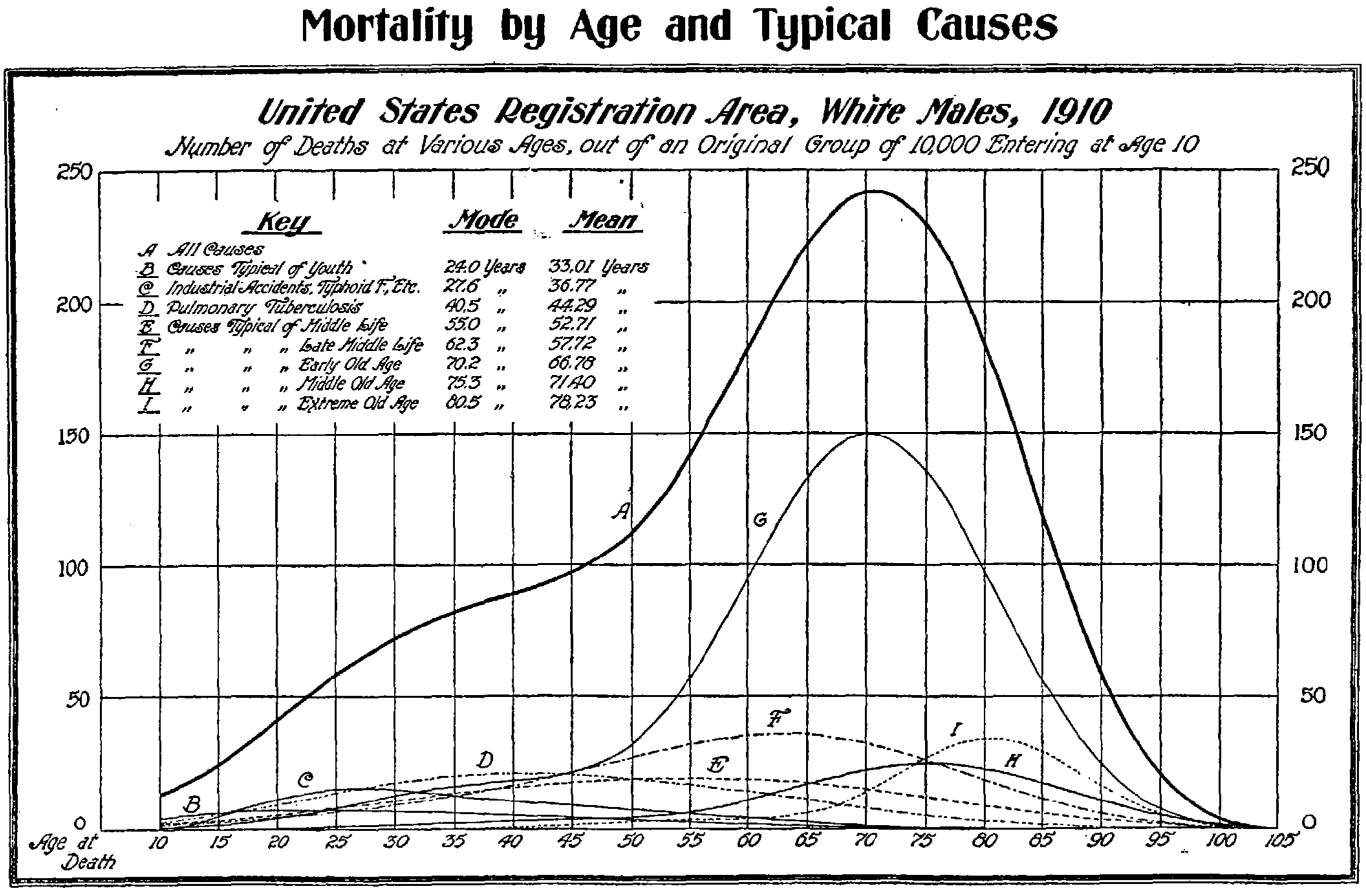
Mortality by Age and Typical Causes
United States Registration Area, White Males, 1910
Number of Deaths at Various Ages, out of an Original Group of 10,000 Entering at Age 10
Key
Mode
Mean
A All Causes
B Causes Typical of Youth 24.0 Years 33.01 Years
C Industrial Accidents, Typhoid F., Etc. 27.6 " 36.77 "
D Pulmonary Tuberculosis 40.5 " 44.29 "
E Causes Typical of Middle Life 55.0 " 52.71 "
F " " " Late Middle Life 62.3 " 57.72 "
G " " " Early Old Age 70.2 " 66.78 "
H " " " Middle Old Age 75.3 " 71.40 "
I " " " Extreme Old Age 80.5 " 78.23 "
250
200
150
100
50
0
Age at Death
10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105
A
B
C
D
E
F
G
H
I

FIG. 1.

For comparative purposes I give below the values of the function q_x alongside the values of the Glover Table for White Males in the Original Registration Area.

TABLE OF VALUES OF $1{,}000q_x$ FOR WHITE MALES CONSTRUCTED FROM THE COMPOUND FREQUENCY CURVE AND CORRESPONDING VALUES OF THE GLOVER WHITE MALES TABLE.

Age.	New Table.	Glover Table
20	4.31	4.89
25	6.17	5.54
30	7.86	6.60
35	9.38	8.52
40	10.71	10.22
45	12.37	12.64
50	15.30	15.53
55	21.09	21.50
60	30.63	30.75
65	44.45	43.79
70	63.11	62.14
75	89.49	92.53
80	118.28	135.75
85	157.96	191.11
90	210.56	255.17
95	282.97	324.86
100	386.71	427.46
105	448.71	582.65

There are, of course, slight discrepancies in the two tables, what indeed might be expected. The Glover Table was constructed on the basis of the estimated population exposed to risk in the *original registration states* and the deaths at various ages for the years 1909, 1910 and 1911, whereas my table was derived from the deaths by cause and age only for the year 1910 for the *total registration area.* A state like Pennsylvania, where for younger ages the accidents from coal mining and other industries are rather high, is not included in the Glover Table. This in connection with the comparatively small area of the southern states included in my table probably accounts for the slight variations. Taken as a whole the two tables show, however, a close agreement, which perhaps may serve as a good test for the validity of the proposed method.

One serious defect I wish, however, to point out in this discussion in connection with the frequency curve of causes of deaths under class C, comprising chiefly deaths arising from occupational hazards. The fit of this particular curve is not very good. It was fitted as a single curve of type B, whereas as a matter of fact it represents a compound curve of two separate curves of which one is a decidedly skew curve of type B and the other a curve of the regular A type. In a table for the total population this defect may not be so serious since the total area of this particular curve is relatively small as compared with the curve areas of causes of deaths in the other classes. In occupations with a pronounced occupational hazard the matter is, however, quite different since the area of this particular curve is of considerable importance as compared with the areas of the other curves, and the total effect in the compound curve tends to give too low mortality in the ages from 15 to 24 and too high a mortality in the ages from 25 to 32. These remarks must be borne in mind in connection with the discussion immediately following.

I shall now, furthermore, as an additional illustration of the method, show how we may construct an occupational mortality table for so distinctive an occupation as that of locomotive engineers. The statistical data forming the basis of this table have most kindly been put at my disposal through the courtesy of Mr. F. S. Crum, assistant statistician of the Prudential Insurance Company of America, who secured the records of deaths by age and cause from more than 7,000 death claims of the "Locomotive Engineers' Mutual Life and Accident Insurance Association" as published in the *Locomotive Engineers' Journal* for the years 1907–1916.

Through a test by correlation methods I found that the parameters to all practical purposes remained essentially unaltered in the various groups, except possibly in groups F and C. The areas of the component curves are, however, quite different from the component areas in the general population as shown by the following table.

TABLE SHOWING RATIO OF AREAS OF COMPONENT CURVES OF LOCOMOTIVE ENGINEERS TO THE CORRESPONDING COMPONENT AREAS OF THE 1910 WHITE MALES IN THE REGISTRATION AREA.

Groups.	Ratio.	Groups.	Ratio.
B	0.2495	F	0.7834
C	3.7988	G	1.0345
D	0.2628	H	0.9639
E	0.7150	I	0.9642

Mortality of Locomotive Engineers. 1907–1916

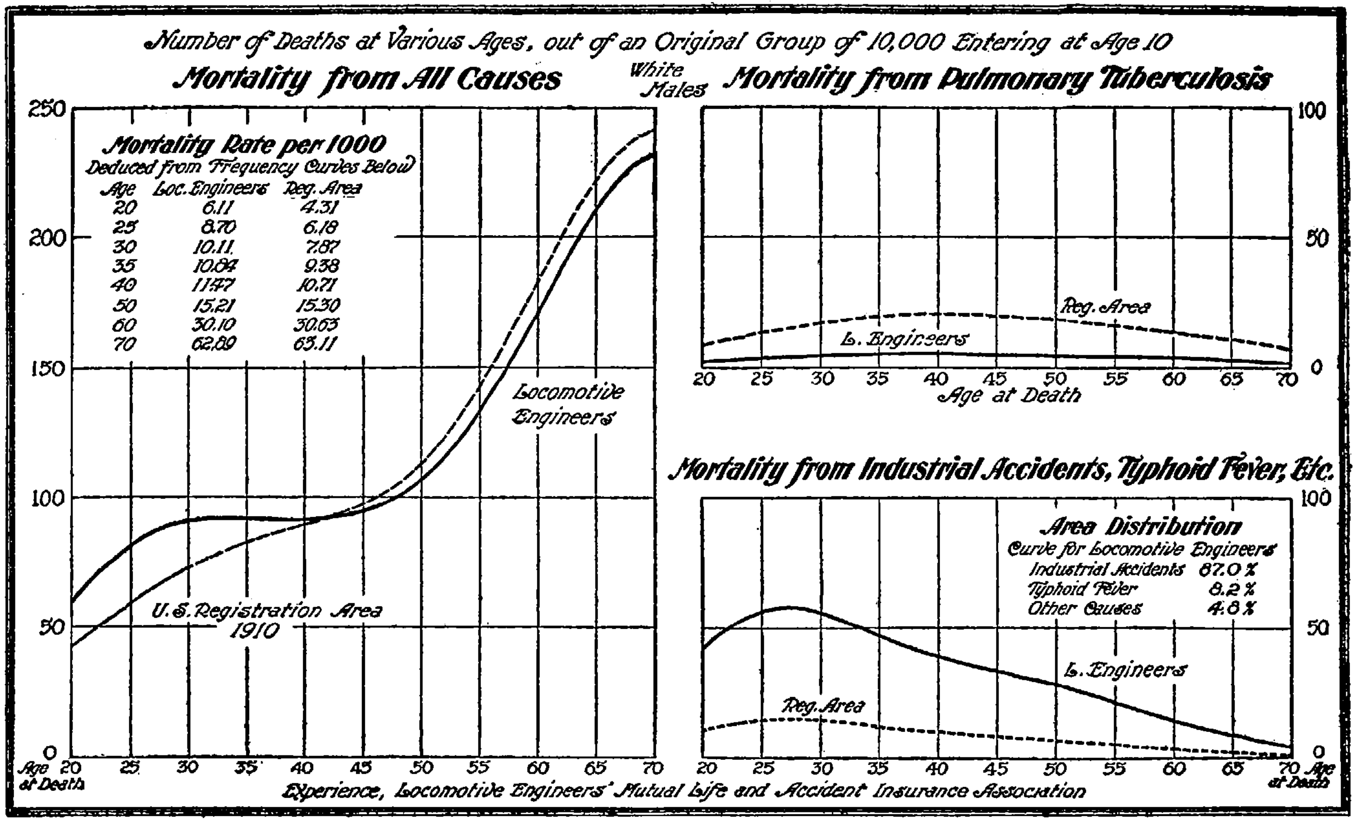

FIG. 2.

This shows that the area of group C, representing deaths chiefly from industrial accidents and typhoid fever, is about 3.8 as great as the corresponding area in the general population, whereas tuberculosis, chiefly represented by curve D, is only about one fourth of that of the general population. This relation is shown graphically in Fig. 2. The same figure shows also the compound frequency curves for locomotive engineers as compared with the general population in the age interval from 20 to 70. The total area of the compound curves for the full span of life must of course be the same for both tables, although differing greatly in a specific age interval. From the compound curve for locomotive engineers I constructed the values of q_x, of which I submit a table for quinquennial ages.

VALUES OF $1000q_x$: LOCOMOTIVE ENGINEERS.

Age: x.	$1000q_x$.	Age: x.	$1000q_x$
20	7.02	50	16.14
25	7.98	55	21.76
30	8.59	60	31.47
35	10.01	65	44.10
40	11.47	70	61.97*
45	13.20		

ADDENDA.

Since the writing of the above article I have made a further study of the stability of the various groups of causes of death. As mentioned previously it is desirable to break up group G in at least two groups. Moreover, by making extended tests for stability, it is of course possible to reach more desirable combinations than the ones originally chosen for the eight groups.

Although the various tests by means of the Lexian-Charlier dispersion theory are not fully completed, I take the liberty to submit the following preliminary scheme of 10 groups, together with the numerical values of the various statistical parameters:

* I wish here to remark that the defect in the frequency curve C in all probability has caused a rise in the mortality rates from ages 25 to 40, while the mortality is too low in the ages below 25. In order to emphasize this I have purposely retained the figures in this paper in order to compare them with a later, and I hope more successful, analysis of the same data given below.

Group.	Mean.	Dispersion.	Skewness.	Excess.
A	(All causes in the compound curve)			
B	Fitted by means of Poisson-Charlier B curves for 5 year intervals			$\lambda = 1.5$
C				$\lambda = 2.1$
D	47.21	14.935	−0.0554	−0.0129
E	40.20	16.359	−0.0406	−0.0286
F	52.28	19.602	+0.0219	−0.0332
G	59.19	18.288	+0.0806	−0.0176
H	64.21	14.578	+0.1193	+0.0256
I	66.93	13.739	+0.1140	+0.0373
J	71.87	13.251	+0.1522	+0.0160
K	77.80	10.328	+0.1408	+0.0488

As an additional illustration of the method I shall try to regraduate the United States Life Table for all Males in the Original Registration States as constructed by Professor James W. Glover from the census of 1910 and the deaths in 1911, 1910 and 1909. We have in this case simply to break up the d_x column of the Glover Table into 10 separate columns as represented by the proportionate mortality of the above groups.

As a matter of illustration let us take the age interval 50–54. According to the census reports the total number of deaths from all causes during 1909, 1910 and 1911 in this interval amounted to 71,252, distributed according to our scheme of classification of causes of death as follows:

Class B	460
C, D	4379
E	11110
F	9197
G	7827
H	20688
I	14130
J	2098
K	1352
	71252

The total number of deaths according to Glover in the same age interval and based upon a radix of 1,000,000 survivors at age 10 is 63,361. Assuming that the proportionate mortality in the Glover Table is the same as in the general population, an assumption which probably is justified, we obtain the following pro rata distribution according to the various classes of deaths:

Class B	418
C, D	3897
E	9878
F	8180
G	6963
H	18394
I	12564
J	1863
K	1204
	63361

Continuing in the same manner for the other age intervals we get a series of columns representing the deaths at various ages in the 10 classes from B to K. Summing up these columns we obtain the area of the various curves expressed in the total number of deaths from the specific groups for all ages. These numbers are then simply multiplied with the equations of the frequency curves for the respective groups of the above classification. The final result is shown in the appended table. (Table I.)

It will now be of interest to make a comparison between the values of q_x as constructed by means of frequency curves and the original values in the Glover Table.

Age.	Glover's q_x.	Fisher's q_x.
10	2.42	1.26
15	2.91	2.95
20	5.03	4.41
25	5.71	5.49
30	6.81	6.81
35	8.74	8.47
40	10.46	10.31
45	12.84	12.72
50	15.81	16.58
55	21.78	22.85
60	31.04	32.40
65	44.06	45.51
70	62.40	62.75
75	92.72	85.64
80	135.64	115.01
85	190.94	151.21
90	253.85	196.12
95	321.76	258.96
100	407.20	359.02
105	526.33	527.27

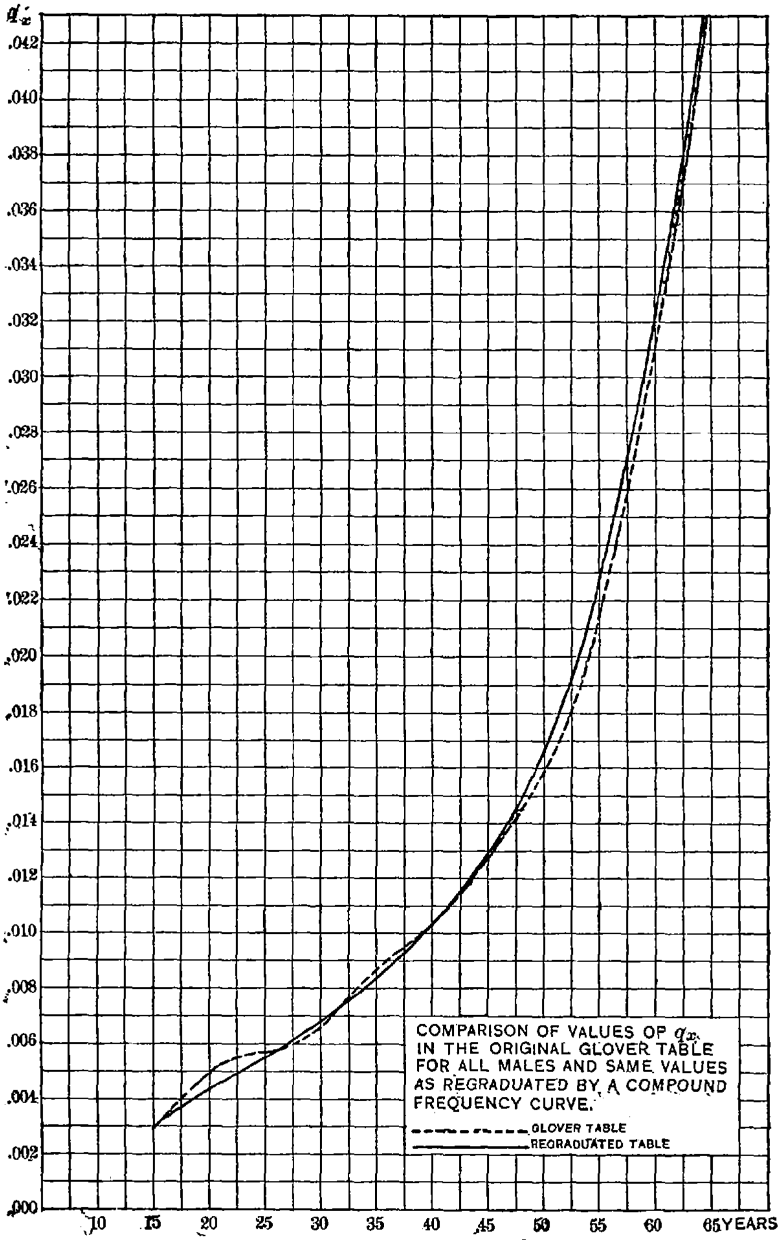
q'_x
.042
.040
.038
.036
.034
.032
.030
.028
.026
.024
.022
.020
.018
.016
.014
.012
.010
.008
.006
.004
.002
.000
10 15 20 25 30 35 40 45 50 55 60 65 YEARS
COMPARISON OF VALUES OF q_x
IN THE ORIGINAL GLOVER TABLE
FOR ALL MALES AND SAME VALUES
AS REGRADUATED BY A COMPOUND
FREQUENCY CURVE.
GLOVER TABLE
REGRADUATED TABLE

FIG. 3.

The larger discrepancies are all in the older ages and in the ages 10–12. The deficiency in the younger ages arises from the fact that the truncated tail end of the B curve of infant mortality reaches into that age period. This curve is not included for ages below 10. The truncated tail end was fitted as a B curve, which gave too low values for the ages 10, 11, 12 and 13 and slightly too high values for the remaining ages. This deficiency is, however, of no practical importance, because of the fact that this particular curve plays no part whatsoever in the ages above 20.

The discrepancies for older ages are at a first sight of a more serious nature. It must, however, be borne in mind that the mortality table for ages beyond 80 is liable to violent perturbations, and no mortality table can lay claim to being exact for those higher ages. Moreover, the frequency curve for extreme old age includes deaths, which in the census reports are classified as due to senility. The number from this cause is quite large. Now whenever senility is given as the cause of death that is probably due to a defective medical diagnosis of the cause of death, which probably was due to a well-defined disease, such as Bright's disease, heart disease, apoplexy, etc. This fact will of course influence the area of the curves including such causes of death, which wrongly are termed senility. I am at present working on the British mortality statistics and hope from the facts contained therein to throw further light on this question.

For practical purposes and rate-making the fit of the various values of q_x for higher ages is of less importance, and the fact that construction of the q_x by means of compound frequency curves for the ages between 15 and 75 shows a remarkable close fit to the original Glover figures seems to me but further evidence of the soundness of the method.*

* As another example of the regraduation of a mortality table we may mention the graduation of the American Experience Table of Mortality by Mr. Arthur Hunter by the Makeham hypothesis. This regraduation is of course nothing more than the fitting of precision observations, assuming the original values of the American Table to be the original observed values of the unknown. Nevertheless Mr. Hunter's table shows for younger ages as large percentage deviations from the original values as the present method of frequency curves, and when it is subjected to the test of stability by dispersion methods the Hunter graduation can by no means be said to be beyond the pale of criticism.

Mortality by Age and Typical Causes

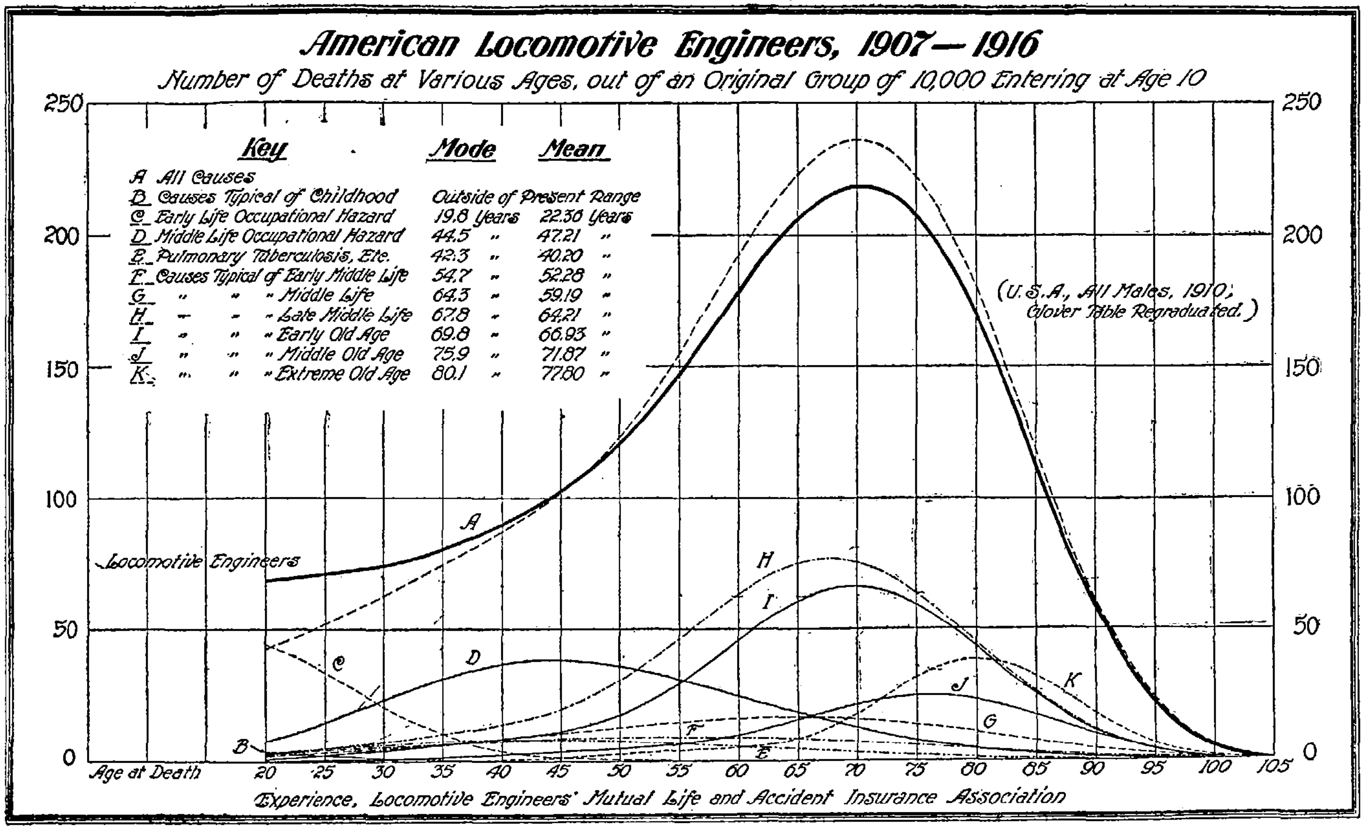

FIG. 4.

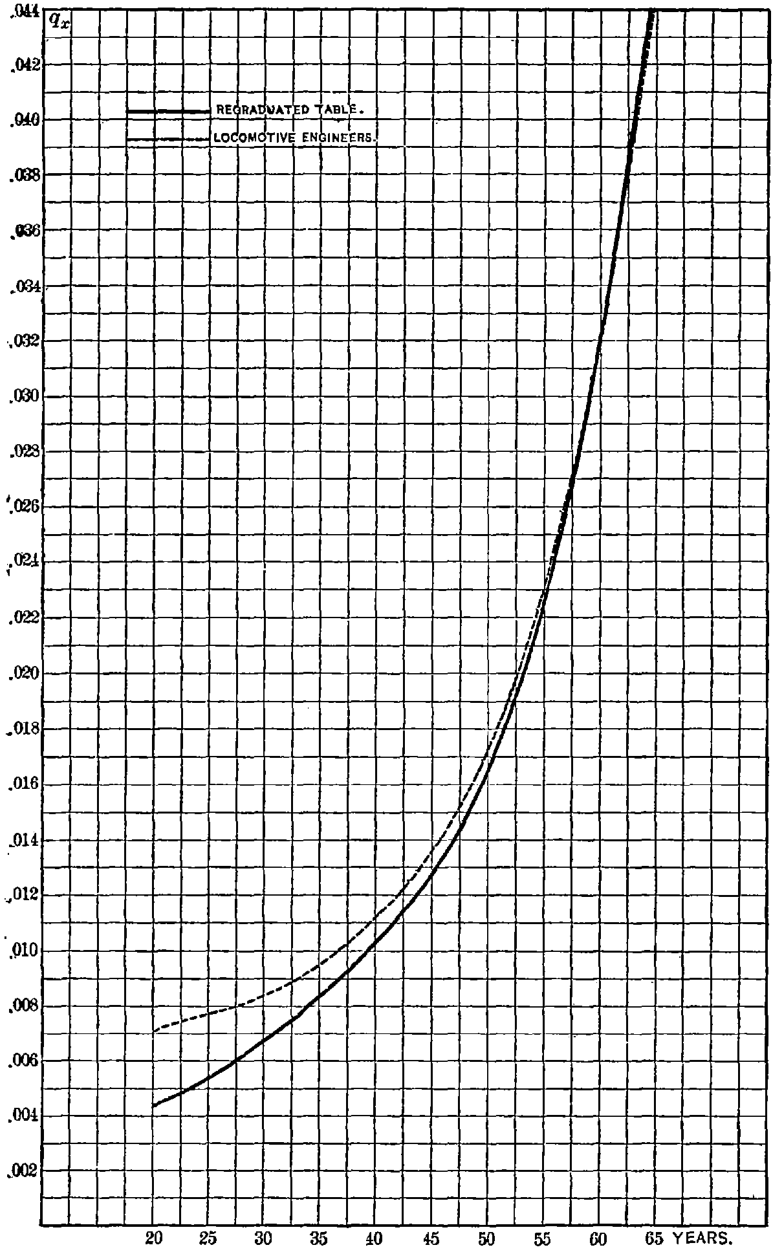

FIG. 5.

Value of q_x as computed from compound frequency curves, showing comparison between the regraduated Glover Table and Locomotive Engineers.

At any rate I trust by a further study of the excellent data contained in the British Register General's Report and additional Dutch, French, Swiss and Scandinavian data to be able to perfect the method for the component skew frequency curves typical of extreme old age.

The defect in the construction of a mortality table for Locomotive Engineers by the aid of the former method of 8 component curves was discussed above. I shall here give the results from an application of the 10 curves where the primarily occupational hazards are represented by curves C and D. The areas of the curves were determined by the usual methods of least squares, and the final result is shown in Table II. The same result is shown graphically in Fig. 4 together with the compound curve of the regraduated Glover Table. This table and the accompanying graph are presented without further comment as to possible merits or defects. I shall instead give a detailed comparison between the observed values of death records by age and cause of Locomotive Engineers and the corresponding values as obtained by frequency curves. Such a comparison is made possible by the following simple method. Choosing at random a certain age interval, say from 55–59, we find in the Locomotive Engineers' Table altogether 80,131 deaths distributed according to the frequency curves as follows:

Class B	0	%
C, D	14216	17.74
E	3126	3.90
F	4467	5.57
G	7729	9.65
H	26620	33.22
I	17855	22.28
J	3876	4.84
K	2242	2.80
Total	80131	100.00

The actual number of observed deaths in this interval was 870. A pro rata distribution according to the percentages above gives the following number in the 10 classes:

Class B	0
C, D	154
E	34
F	48
G	84
H	289
I	194
J	42
K	25
Total	870

The distribution actually observed according to the same classes of causes of deaths was as follows:

Class B	1
C, D	134
E	31
F	34
G	90
H	313
I	194
J	44
K	29
Total	870

A similar tabular arrangement is shown below for all age groups of the calculated (expected) and observed number of deaths in the various classes of deaths.

An actual test for the "goodness of fit" would take considerable time, and I leave it to my readers to form their own judgment as to the actual success of the construction on the basis of the above tabular arrangement of observed and expected numbers of deaths.

The mortality tables as constructed by means of a few simple theorems from the mathematical theory of probabilities are here viewed in their actuarial aspect only. Much more interesting and weightier questions of a purely philosophical and biometric nature in connection with the study of human mortality must unfortunately be left out in this short article. The final values are the results of a purely deductive analysis by means of mathematical methods. *We must, however, not forget that the deduction was founded upon a hypothesis as propounded at the beginning of this note.* If it can be proven that this hypothesis is wrong, none of the

OBSERVED AND CALCULATED (EXPECTED) DEATHS ACCORDING TO TYPICAL CAUSES. (LOCOMOTIVE ENGINEERS).

	20–24.		25–29.		30–34.		35–39.		40–44.		45–49.	
	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.
K	—	—	0	0	2	0	1	2	8	8	9	17
J	—	1	1	1	12	8	26	19	26	29	37	37
I	—	—	6	9	26	38	64	70	97	95	142	125
H	1	1	9	16	36	64	100	114	166	161	239	222
G	2	1	11	11	48	41	65	73	85	95	80	108
F	—	1	18	14	43	46	63	70	80	78	70	80
E	1	1	21	15	60	52	81	77	83	82	67	74
C&D	16	14	131	128	344	321	431	405	410	409	379	361
B	0	1	1	4	4	5	1	2	2	0	1	0
	20		198		575		832		957		1024	

	50–54.		55–59.		60–64.		65–69.		70–74.		75–79.		80 and Over.	
	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.
K	23	26	29	24	19	19	30	26	24	27	25	25	40	32
J	46	43	44	42	37	40	37	39	28	27	12	18	13	18
I	168	172	194	194	177	169	139	137	80	76	44	38	26	27
H	297	289	313	289	226	223	159	164	78	84	35	41	16	27
G	109	108	90	85	48	53	42	35	25	18	18	11	17	9
F	74	70	34	48	36	28	16	17	7	9	4	4	5	4
E	51	57	31	34	18	17	5	8	7	3	2	1	1	1
C&D	265	270	134	154	59	71	31	34	7	12	1	3	2	2
B	2	0	1	0										
	1035		870		620		459		256		141		120	

	All Ages.				Obs.-Calc.	Diff. in Per Cent. of Obs.
	Obs.	Per Cent.	Calc.	Per Cent.		
K	210	2.95	206	2.90	+ 4	1.90
J	319	4.49	321	4.52	− 2	0.62
I	1163	16.36	1150	16.18	+ 13	1.12
H	1675	23.56	1695	23.85	− 20	1.19
G	640	9.02	648	9.12	− 8	1.25
F	450	6.33	469	6.59	− 19	4.22
E	428	6.02	422	5.94	+ 6	1.39
C & D	2210	31.10	2184	30.73	+ 26	1.18
B	12	0.17	12	0.17	+ 0	0.00
Total...........	7107	100.00	7107	100.00		

results will be valid in general. On the other hand, the progress of the most exact sciences as those of physics and chemistry has shown that deductive mathematical reasoning based upon a sound and true hypothesis leads to more general and trustworthy results than the ones obtained by means of purely empirical methods, which in my opinion have been employed by far too great an extent by actuaries and statisticians, and unfortunately at the loss of sound logic and common sense reasoning. Too many actuaries labor still under the fallacious ideas of John Stuart Mill that it is possible to deduce the corpus of human knowledge by what he termed an "*inductio per simplicem enumerationem.*"

In the meantime I beg my readers to remember and bear in mind the very scant and imperfect data I have had at my disposal and which nevertheless by a purely mathematical analysis have yielded plausible results, which certainly never could have been obtained by ordinary actuarial methods. If American actuaries in the future should turn their attention to the application of frequency curves and modern statistical methods, I can only wish that the above remarks may prove a modest step in the final solution of the problem of the construction of mortality tables along purely deductive mathematical lines.

Considering the fact, however, that I have had at my disposal death records only and not the slightest clue as to the age distribution of the numbers exposed to risk at various ages, I may perhaps be justified in advocating the use of modern statistical-mathematical methods in such actuarial practice, where the conventional methods leave us at a complete loss as to actual results. Moreover, I may in conclusion perhaps be allowed to assert that the results I have obtained from the use of mathematical-statistical methods do at least entitle such methods to a fair and impartial consideration instead of being looked upon with distrust and as having a theoretical interest only.

As a conclusion to these brief remarks I take great pleasure in acknowledging my thanks to a number of friends and statistical confrères who have given me valuable help in this little study on mortality. To Mr. F. L. Hoffman, statistician of the Prudential Insurance Company of America, and to Mr. F. S. Crum, assistant statistician of the same company, my most grateful thanks are due for the generosity with which they have put the necessary statistical material at my disposal. My thanks are also due to Mr. Knud Stoumann, formerly of the government statistical bureau of Copenhagen and now connected with the Prudential, who has rendered valuable assistance in the classification of the various causes of death, and to Mr. E. E. A. Fisher and Mr. C. Balck, also of the Prudential staff, for drawing the figures and making most of the extended numerical calculations in the actual construction of the tables.

TABLE I.

REGRADUATED LIFE TABLE FOR MALES IN THE ORIGINAL REGISTRATION STATES: 1910, SHOWING COMPOUND FREQUENCY CURVE OF NUMBER OF DEATHS BY VARIOUS AGES AND ITS 10 COMPONENT CURVES OF NUMBER OF DEATHS BY AGE AND TYPICAL CAUSES.

Age.	K.	J.	I.	H.	G.	F.	E.	D.	C.	B.	A, d_x.	l_x.	$1{,}000q_x$.
10			11	56	91	231	259	22	241	345	1,256	1,000,000	1.26
11			22	68	104	259	309	31	418	386	1,597	998,744	1.60
12			31	83	116	295	364	40	577	459	1,965	997,147	1.97
13			41	96	129	329	425	52	697	517	2,286	995,182	2.30
14			50	116	146	363	490	66	816	575	2,622	992,896	2.64
15			59	132	761	400	573	81	896	624	2,926	990,274	2.95
16			74	154	181	433	650	99	975	656	3,222	987,348	3.26
17			89	181	199	479	732	121	1,035	692	3,528	984,126	3.58
18			107	207	222	521	819	141	1,075	725	3,817	980,598	3.89
19		6	126	240	243	564	909	168	1,095	710	4,061	976,781	4.16
20		8	150	270	268	610	1,003	197	1,115	667	4,288	972,720	4.41
21		10	177	310	291	656	1,100	229	1,075	631	4,479	968,432	4.63
22		12	205	346	320	714	1,200	257	1,035	577	4,666	963,953	4.84
23		15	237	379	346	762	1,299	295	975	519	4,827	959,287	5.03
24		19	272	426	379	812	1,401	334	915	464	5,022	954,460	5.26
25		23	312	474	406	861	1,519	369	836	408	5,208	949,438	5.49
26		28	353	520	442	912	1,619	411	776	352	5,413	944,230	5.73
27		33	397	570	473	961	1,718	455	717	302	5,626	938,817	5.99
28		40	449	620	505	1,012	1,813	500	657	255	5,851	933,191	6.27
29		46	493	664	544	1,061	1,904	539	577	212	6,040	927,340	6.51
30		55	543	715	577	1,111	1,993	585	517	174	6,270	921,300	6.81
31		65	593	770	620	1,159	2,074	630	458	143	6,512	915,030	7.12
32		73	649	823	656	1,216	2,152	668	398	116	6,751	908,518	7.43
33		85	700	879	700	1,262	2,222	711	358	94	7,011	901,767	7.77
34		96	752	936	738	1,307	2,297	752	299	80	7,257	894,756	8.11
35	6	109	808	994	785	1,350	2,353	784	259	65	7,513	884,499	8.47
36	14	122	858	1,045	825	1,392	2,400	821	219	49	7,745	879,986	8.80
37	20	137	907	1,108	874	1,432	2,443	853	179	40	7,993	872,241	9.16
38	27	151	957	1,177	915	1,470	2,476	878	139	28	8,218	864,248	9.51
39	35	166	1,017	1,250	966	1,508	2,501	904	119	22	8,488	856,030	9.92
40	44	183	1,070	1,330	1,010	1,543	2,519	925	100	15	8,739	847,542	10.31
41	57	197	1,127	1,423	1,050	1,575	2,530	940	80	11	8,990	838,803	10.72
42	72	213	1,198	1,527	1,107	1,613	2,535	952	60	9	9,286	829,813	11.19
43	89	228	1,270	1,625	1,162	1,641	2,531	961	60	7	9,564	820,527	11.66
44	109	244	1,347	1,755	1,207	1,668	2,521	964	40	4	9,859	810,963	12.16
45	130	258	1,442	1,903	1,263	1,692	2,506	964	40	2	1,0190	801,104	12.72
46	157	275	1,567	2,070	1,308	1,714	2,484	959	40	1	1,0575	790,914	13.37
47	185	290	1,695	2,260	1,355	1,735	2,456	951	20		1,0947	780,334	14.03
48	215	307	1,844	2,470	1,410	1,753	2,423	939	20		1,1381	769,392	14.79
49	243	323	2,018	2,703	1,456	1,769	2,387	923			1,1822	758,011	15.60
50	276	343	2,249	2,958	1,512	1,783	2,346	904			1,2371	746,189	16.58
51	307	364	2,482	3,193	1,555	1,795	2,301	885			1,2882	733,818	17.55
52	335	385	2,743	3,486	1,608	1,803	2,244	860			1,3464	720,936	18.68
53	365	413	3,075	3,799	1,650	1,810	2,190	833			1,4135	707,472	19.98
54	391	440	3,395	4,128	1,699	1,813	2,134	808			1,4808	693,337	21.36
65	412	476	3,743	4,465	1,738	1,814	2,076	778			1,5502	678,529	22.85
56	433	512	4,116	4,792	1,782	1,811	2,014	746			1,6206	663,027	24.44
57	451	560	4,565	5,155	1,815	1,806	1,949	711			1,7012	646,821	26.30
58	468	607	4,973	5,450	1,853	1,798	1,882	685			1,7716	629,809	28.13
59	488	667	5,389	5,783	1,877	1,787	1,813	650			1,8454	612,093	30.15

TABLE I (*continued*).

REGRADUATED LIFE TABLE FOR MALES IN THE ORIGINAL REGISTRATION STATES: 1910, SHOWING COMPOUND FREQUENCY CURVE OF NUMBER OF DEATHS BY VARIOUS AGES AND ITS 10 COMPONENT CURVES OF NUMBER OF DEATHS BY AGE AND TYPICAL CAUSES.

Age.	K.	J.	I.	H.	G.	F.	E.	D.	C.	B.	A, d_x.	l_x.	$1{,}000q_x$.
60	513	726	5,866	6,100	1,900	1,773	1,742	616			19,236	593,639	32.40
61	543	800	6,276	6,395	1,923	1,755	1,657	587			19,936	574,403	34.71
62	587	883	6,674	6,664	1,937	1,730	1,582	553			20,610	554,467	37.17
63	649	956	7,047	6,889	1,947	1,706	1,507	519			21,220	533,857	39.75
64	725	1,046	7,438	7,076	1,952	1,679	1,431	490			21,837	512,637	42.60
65	832	1,128	7,740	7,227	1,949	1,648	1,354	457			22,335	490,800	45.51
66	969	1,222	7,999	7,328	1,942	1,614	1,277	425			22,776	468,465	48.62
67	1,120	1,305	8,210	7,386	1,927	1,577	1,200	394			23,119	445,689	51.87
68	1,318	1,397	8,358	7,394	1,908	1,537	1,123	368			23,403	422,570	55.38
69	1,546	1,474	8,442	7,346	1,879	1,494	1,035	339			23,553	399,167	59.01
70	1,739	1,556	8,457	7,248	1,850	1,448	962	311			23,571	375,614	62.75
71	2,049	1,625	8,400	7,099	1,808	1,390	889	288			23,548	352,043	66.89
72	2,338	1,683	8,280	6,934	1,768	1,340	819	262			23,424	328,495	71.31
73	2,630	1,730	8,102	6,682	1,713	1,287	750	238			23,132	305,071	75.82
74	2,885	1,765	7,869	6,422	1,664	1,233	685	218			22,741	281,939	80.66
75	3,154	1,787	7,539	6,114	1,609	1,177	622	196			22,198	259,198	85.64
76	3,385	1,792	7,202	5,776	1,540	1,120	562	176			21,553	237,000	90.94
77	3,576	1,783	6,825	5,412	1,479	1,062	507	160			20,804	215,447	96.56
78	3,724	1,760	6,362	5,034	1,404	1,004	444	142			19,874	194,643	102.10
79	3,820	1,720	5,929	4,700	1,337	946	396	125			18,973	174,769	108.56
80	3,852	1,670	5,482	4,310	1,256	887	350	111			17,918	155,796	115.01
81	3,829	1,602	5,029	3,916	1,188	819	309	98			16,790	137,878	121.77
82	3,744	1,530	4,512	3,531	1,104	762	270	86			15,539	121,088	128.33
83	3,614	1,440	4,064	3,156	1,035	706	235	74			14,324	105,549	135.71
84	3,423	1,340	3,632	2,795	952	652	203	66			13,063	91,225	143.20
85	3,186	1,248	3,216	2,454	884	600	174	57			11,819	78,162	151.21
86	2,944	1,138	2,772	2,178	804	548	148	48			10,680	66,343	159.47
87	2,654	1,040	2,410	1,880	740	500	122	42			9,388	55,763	168.36
88	2,347	928	2,074	1,605	664	453	102	36			8,209	46,374	177.01
89	2,071	831	1,730	1,353	605	409	85	30			7,114	38,166	186.40
90	1,770	725	1,458	1,128	547	367	70	25			6,090	31,052	196.12
91	1,482	636	1,218	931	482	320	58	21			6,148	24,962	206.23
92	1,272	540	1,005	757	430	283	46	18			4,341	19,814	219.59
93	1,002	462	797	624	374	249	37	15			3,560	15,463	230.23
94	790	380	643	493	329	218	29	12			2,894	11,903	243.13
95	623	316	512	383	279	188	22	10			2,333	9,009	258.96
96	464	250	390	288	242	162	16	8			1,825	6,676	273.37
97	342	193	299	210	201	138	10	7			1,400	4,851	288.60
98	240	150	229	149	170	116	9	5			1,068	3,451	309.48
99	171	108	170	99	132	97	6	4			787	2,383	330.26
100	109	78	120	65	113	80	4	4			573	1,596	359.02
101	67	49	87	35	87	62	1				388	1,023	379.28
102	40	28	72	12	68	49	1				270	635	425.20
103	24	10	41		49	37					161	365	441.10
104	7		26		34	27					94	204	217.02
105			18		21	19					58	110	527.27
106			13		11	12					36	52	692.31
107			7		3	6					16	16	1,000.00

TABLE II

Mortality Table for American Locomotive Engineers.

Constructed by means of compound frequency curve method from records of deaths by cause and age of the Locomotive Engineer's Mutual Life and Accident Association (1907–1916).

Age.	*K.*	*J.*	*I.*	*H.*	*G.*	*F.*	*E.*	*D.*	*C.*	*B.*	$A=d_x$.	l_x.	$1000q_x$.
10			9	58	78	114	83	88	960	183	1,573	1,000,000	1.57
11			17	71	89	128	99	123	1,665	205	2,397	998,427	2.40
12			24	86	99	146	117	159	2,298	243	3,172	996,030	3.18
13			32	100	110	163	136	207	2,776	274	3,798	992,858	3.83
14			39	120	124	180	157	263	3,249	305	4,437	989,060	4.49
15			46	137	137	198	184	323	3,568	331	4,924	984,623	5.00
16			58	160	154	214	209	394	3,882	348	5,419	979,699	5.53
17			70	188	170	237	235	482	4,121	367	5,870	974,280	6.02
18			84	215	189	258	263	561	4,281	385	6,236	968,410	6.44
19		8	99	249	207	279	292	669	4,360	377	6,540	962,174	6.80
20		11	118	280	229	302	322	784	4,440	354	6,840	955,634	7.16
1		14	139	322	248	325	353	912	4,281	335	6,929	948,794	7.30
2		16	161	359	273	354	385	1,023	4,121	306	6,998	941,865	7.43
3		21	186	393	295	377	417	1,175	3,883	275	7,022	934,867	7.51
4		26	213	442	323	402	450	1,330	3,613	246	7,045	927,845	7.59
5		32	245	492	346	426	488	1,469	3,359	216	7,073	920,800	7.68
6		38	277	540	377	452	520	1,637	3,090	187	7,118	913,727	7.79
7		45	311	592	403	476	552	1,812	2,855	160	7,206	906,609	7.95
8		55	352	644	431	501	582	1,991	2,616	135	7,307	899,403	8.12
9		63	386	689	464	526	611	2,146	2,378	112	7,375	892,096	8.27
30		76	426	742	492	550	640	2,329	2,059	93	7,407	884,721	8.37
1		89	465	799	529	574	666	2,509	1,824	76	7,531	877,314	8.58
2		100	509	854	559	602	691	2,660	1,585	62	7,622	869,783	8.76
3		117	549	912	597	625	713	2,831	1,345	50	7,739	862,161	8.98
4		132	589	972	629	647	738	2,994	1,191	42	7,934	854,422	9.29
5	6	150	633	1,032	669	669	755	3,122	1,031	34	8,101	846,488	9.57
6	14	168	672	1,085	703	689	771	3,269	872	26	8,269	838,387	9.86
7	20	188	711	1,150	745	709	784	3,397	713	21	8,438	830,118	10.16
8	27	207	750	1,222	780	728	795	3,496	554	15	8,574	821,680	10.43
9	35	228	797	1,298	824	747	803	3,600	474	12	8,818	813,106	10.84
40	44	251	838	1,380	861	764	809	3,683	398	8	9,036	804,288	11.23
1	57	271	883	1,477	895	780	812	3,743	318	6	9,242	795,252	11.62
2	72	293	939	1,585	944	799	814	3,791	239	5	9,481	786,010	12.06
3	89	313	995	1,687	982	813	813	3,827	239	3	9,761	776,529	12.57
4	109	335	1,055	1,822	1,029	826	809	3,839	159	2	9,985	766,768	13.02
5	130	354	1,130	1,975	1,068	838	805	3,839	159	1	10,299	756,783	13.61
6	156	378	1,228	2,149	1,115	849	798	3,819	159		10,651	746,484	14.27
7	184	398	1,328	2,346	1,155	859	789	3,787	80		10,926	735,833	14.85
8	214	422	1,445	2,564	1,202	868	778	3,739	80		11,312	724,907	15.60
9	242	444	1,581	2,806	1,242	876	766	3,675			11,632	713,595	16.30
50	275	471	1,762	3,071	1,289	883	753	3,600			12,104	701,963	17.24
1	306	500	1,945	3,315	1,326	889	739	3,524			12,544	689,859	18.18
2	334	529	2,149	3,619	1,371	893	721	3,425			13,041	677,315	19.25
3	364	568	2,409	3,943	1,407	896	703	3,317			13,607	664,274	20.48
4	390	605	2,660	4,285	1,449	898	685	3,218			14,190	650,667	21.81
5	410	654	2,933	4,635	1,482	899	667	3,098			14,778	636,477	23.22
6	431	703	3,225	4,974	1,519	897	647	2,971			15,367	621,699	24.72
7	449	769	3,577	5,351	1,548	895	626	2,831			16,046	606,332	26.46
8	466	834	3,897	5,657	1,580	891	604	2,728			16,657	590,286	28.22
9	486	916	4,223	6,003	1,600	885	582	2,588			17,283	573,629	30.13
60	511	997	4,597	6,332	1,620	878	559	2,453			17,947	556,346	32.26
1	541	1,099	4,918	6,638	1,640	869	532	2,337			18,574	538,399	34.50
2	585	1,213	5,229	6,917	1,652	857	508	2,202			19,163	519,825	3.668

Age.	K.	J.	I.	H.	G.	F.	E.	D.	C.	B.	$A=d_x$.	l_x.	$1000q_x$.
63	647	1,314	5,522	7,151	1,660	845	484	2,067			19,690	500,662	39.33
4	722	1,437	5,829	7,345	1,664	832	459	1,951			20,239	480,972	42.08
5	829	1,550	6,065	7,502	1,662	816	435	1,820			20,679	460,733	44.88
6	965	1,679	6,268	7,607	1,656	799	410	1,692			21,076	440,054	47.89
7	1,116	1,793	6,434	7,667	1,643	781	385	1,569			21,388	418,978	51.05
8	1,313	1,919	6,550	7,675	1,627	761	361	1,465			21,671	397,590	54.51
9	1,540	2,025	6,615	7,625	1,602	740	332	1,350			21,829	375,919	58.07
70	1,732	2,138	6,627	7,524	1,577	717	309	1,238			21,862	354,090	61.74
1	2,041	2,233	6,583	7,369	1,542	689	285	1,147			21,889	332,228	65.89
2	2,329	2,312	6,489	71,98	1,508	664	263	1,043			21,806	310,339	70.27
3	2,620	2,377	6,349	6,936	1,461	638	241	948			21,570	288,533	74.76
4	2,874	2,425	6,167	6,666	1,419	611	220	868			21,250	266,963	79.60
5	3,142	2,455	5,908	6,347	1,372	583	200	780			20,787	245,713	84.59
6	3,372	2,462	5,644	5,996	1,313	555	180	701			20,223	224,926	89.91
7	3,563	2,450	5,349	5,618	1,261	526	163	637			19,567	204,703	95.59
8	3,710	2,418	4,986	5,225	1,197	497	143	565			18,741	185,136	101.23
9	3,806	2,363	4,646	4,879	1,140	469	127	498			17,928	166,395	107.74
80	3,838	2,294	4,296	4,474	1,071	439	112	442			16,966	148,467	116.42
1	3,815	2,201	3,941	4,065	1,013	406	99	390			15,930	131,501	121.14
2	3,730	2,102	3,535	3,665	941	378	87	342			14,780	115,571	127.89
3	3,601	1,979	3,184	3,276	883	350	75	295			13,643	100,791	135.36
4	3,410	1,841	2,846	2,901	812	323	65	263			12,461	87,148	142.99
5	3,174	1,715	2,520	2,547	754	297	56	227			11,290	74,687	151.17
6	2,933	1,564	2,172	2,261	686	272	48	191			10,127	63,397	159.74
7	2,644	1,429	1,888	1,952	631	248	39	167			8,998	53,270	168.91
8	2,338	1,275	1,625	1,666	566	224	33	143			7,870	44,272	177.76
9	2,063	1,142	1,356	1,404	516	203	27	119			6,830	36,402	187.63
90	1,763	996	1,142	1,171	466	182	22	100			5,842	29,572	197.55
1	1,476	874	954	966	411	159	19	84			4,943	23,730	208.30
2	1,267	742	788	786	367	140	15	72			4,177	18,787	222.33
3	998	635	624	648	319	123	12	60			3,419	14,610	234.02
4	787	522	504	512	281	108	9	48			2,771	11,191	247.61
5	621	434	401	398	238	94	7	40			2,233	8,420	265.20
6	467	344	307	299	206	80	5	32			1,740	6,187	281.23
7	341	265	234	218	171	68	3	28			1,328	4,447	298.63
8	239	206	179	155	145	58	3	20			1,005	3,119	322.22
9	170	148	133	103	113	48	4	16			735	2,114	347.68
100	109	107	94	67	96	40		13			526	1,379	381.43
1	67	67	68	36	74	31		4			347	853	406.80
2	40	39	56	12	58	24					229	506	452.57
3	24	14	32		42	18					130	277	469.31
4	7		20		29	13					69	147	469.39
5			14		18	10					42	78	538.46
6			10		9	6					25	36	669.44
7			5		3	3					11	11	1,000.00