COMPILATION AND GRADUATION

of

STATISTICS RELATING TO THE DISABLEMENT

of

LIVES

By

Elgin Robertson Batho
B. A., Univ. of Manitoba, 1925

A Thesis

Submitted to the
University of Manitoba
In Partial Fulfilment
Of the Requirements
for the

MASTER OF ARTS

Degree

April, 1927

COMPILATION AND GRADUATION OF STATISTICS
RELATING TO THE DISABLEMENT OF LIVES

The role of the science of statistics is becoming increasingly important, not only in the academic world, but also in the business world. This is perhaps true in the life insurance business to a greater extent than in other fields. Where insurance companies formerly carried on their operations relying upon the judgment of their officials, they are seeking more and more to base their conclusions upon the deductions to be obtained from statistical data. Such data are now being compiled and used on as nearly a scientific basis as possible.

One of the comparatively new fields into which life insurance companies have entered is that of the granting of certain benefits to their policyholders who pay an extra premium therefor, upon their becoming what may, in general, be termed "totally and permanently disabled". The definition of such disability differs from company to company, and the benefits allowed upon disability also vary. While there are differences between companies, each company does its best to obtain a homogeneous experience by treating all cases that arise in as uniform a manner as possible.

It is not the purpose of this thesis to give a history of the writing of disability benefit insurance, or of the benefits allowed, or premium rates charged, or of the difficulties experienced in this field of insurance, but rather to give illustrations of how statistical data along this line are compiled and graduated, in order to be of the most use to the company.

Specimen Disability Benefit Clause.

Nevertheless, it is essential that something be said about the disability benefit clauses inserted in life insurance contracts. A typical modern clause, issued by one company, reads as follows:

"If, after one year's premiums shall have been paid on this policy and, before default in the payment of any subsequent premiums, the insured shall furnish to the Company due proof that, before attaining the age of sixty, he has become wholly disabled by bodily injury or disease so that he is, and thereby will be permanently and continuously unable to engage in any occupation whatsoever for remuneration or profit, and that such disability has existed continuously not less than sixty days prior to the furnishing of proof, thereupon the Company will grant the following benefits:

A. Waiver of Premium. The Company, by endorsement hereon, shall waive the payment of the premiums which thereafter may become due under this policy during the continuance of the said total disability of the insured. In making any settlement under this policy, the Company shall not deduct any part of the premiums so waived, and the non-forfeiture values of this policy shall increase from year to year in the same manner as though any premium waived under this provision had been paid in cash.

B. Annuity Payment. Furthermore, the Company will pay to the insured a monthly sum equal to one one-hundredth of the face of this policy, the first monthly payment to be made six months after receipt of due proof of the said total disability, accompanied by this policy for endorsement, and subsequent payments monthly thereafter during the continuance of the said total disability of the insured prior to the maturity of this policy. The Company will admit the age of the insured when furnished with satisfactory evidence of the date of birth, and reserves the right to require such proof of date of birth at the time proof of disability is furnished. The amount of the policy, payable at maturity as an endowment or as a death claim, shall not be reduced by any payments made under this disability provision."

(See "Total Permanent Disability Benefits in Relation to Life Insurance", by Arthur Hunter, pages 6 and 7)

In almost all cases, it is stated that the loss of both eyes, or of both hands, or of both feet, or one hand and one foot, shall be considered as total and permanent disability, even though such loss may not prevent the person from carrying on any occupation.

The above disability clause is not the most modern and liberal now in use. A more liberal clause provides that total disability shall be considered to be permanent after it has continuously existed for ninety consecutive days, and that the benefits allowed by the clause shall commence at the end of that period, or one month later.

Principles Governing an Investigation into a Disability Experience.

The policyholders of an insurance company who are not disabled so as to be eligible to receive the disability benefit are termed "active lives". Those who are so disabled are termed "invalid lives", or "disabled lives". Two of the most important factors entering into the determination of the scientifically correct and adequate premiums to be charged by an insurance company for its disability benefits are the rate of disability among those exposed to risk of disability, and the rate of recovery and death among those disabled. "Withdrawals" among disabled lives may be defined as the sum of the deaths and recoveries among such lives. When they have sufficient data, companies are more and more investigating their experience as to their rate of disability and of withdrawal.

In making an investigation to determine the rate of disability, a company includes only those policyholders who have the disability benefit clause attached to their policies. Other policyholders would have no incentive to report their disability, and the company could not obtain accurate data in connection with them. The policyholders who have the disability clause in their policies may be relied upon to report their disability.

An investigation of a disability experience may be made by policies, by lives, or by amounts. An investigation by lives is rather laborious, as it is rather difficult to eliminate all but the first policy on each life. Such a method is also not adapted to the construction of select tables. An investigation by policies is more convenient, and usually satisfactory, but does not give the actual monetary effect of the company's disability experience, as the amounts of the policies vary greatly. An investigation by amounts, although it will quite likely present certain irregularities or distortions, will give this monetary effect in which the company is primarily interested. As far as financial effect is concerned, it makes no difference to the company whether one policy for $10,000, or ten policies for $1000 each, become disability claims. Below are given the details of an investigation into the rate of disability, based upon amounts insured.

If the amount of the data was at all large, it is likely that the statistics would be compiled with the help of perforated cards, and electric sorting and tabulating machines. If these aids were adopted, and if the investigation was based upon amounts, instead of upon lives or policies, a card would be perforated, by means of perforating machines, for each policy included within the investigation, and the amount of insurance, taken to the nearest hundred dollars, would be punched on the card. Illustration number one shows a blank card to be used in this connection.

A great deal of data is included on the disability investigation card. In the Illustration, the amount of insurance on which the disability benefit is based is punched in columns 15, 16, and 17. The card provides for an amount of insurance only up to one hundred thousand dollars. The policy number, which is punched in columns 9 to 14 on the card in the illustration, is necessary for identification purposes. The information punched on the card is obtained from the company records, generally from certain record cards, one of which is kept for each policy. The sex is punched in column 18, the code being, in the case of the card shown in Illustration number 1:

Male	Punch	out	the	"0"	of	column	18
Single Female	"	"	"	"1"	"	"	18
Married Female	"	"	"	"2"	"	"	18

The agencies of the company are numbered and the code number is punched in the next three columns. The year of issue, as indicated in the next two columns, is self-explanatory. The codes for the plan and for the residence are indicated next. The age at issue comes next, and the class, whether standard or substandard, is indicated in column 31.

The occupation, punched in the next two columns, may be used as a basis for a further investigation of the experience in regard to occupations. These occupations are grouped into only a few broad classes. The build of the insured, that is, the relation between his height and weight, is punched on the card, as it is believed that it influences, to a certain extent, the likelihood of his becoming disabled. For example, it has long been recognized that young underweights are much more susceptible to tuberculosis than those who are older and somewhat overweight. Column 36 gives the impairments in the family history of the insured which have a bearing on the risk of disablement. The next four columns indicate the particular form of Double Indemnity and Disability Benefit clauses contained in the policy.

SPECIMEN CARD USED IN COMPILING THE DATA

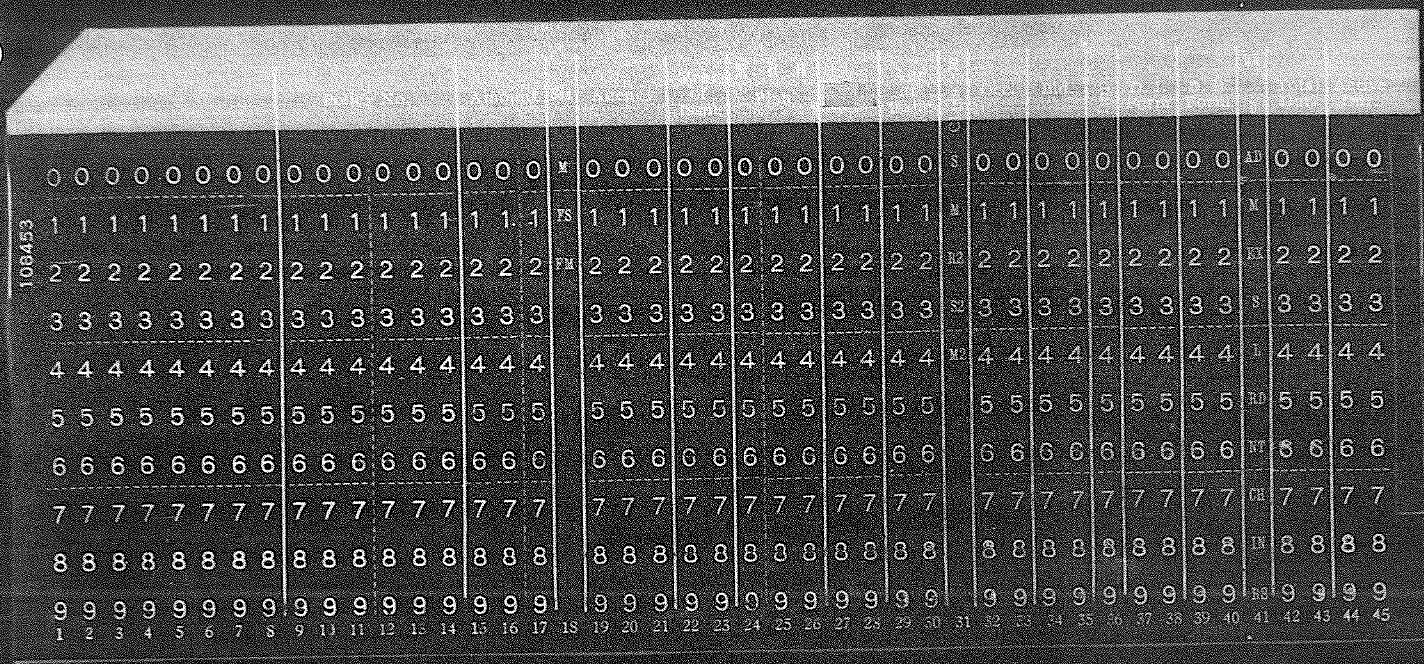

Illustration Number 1

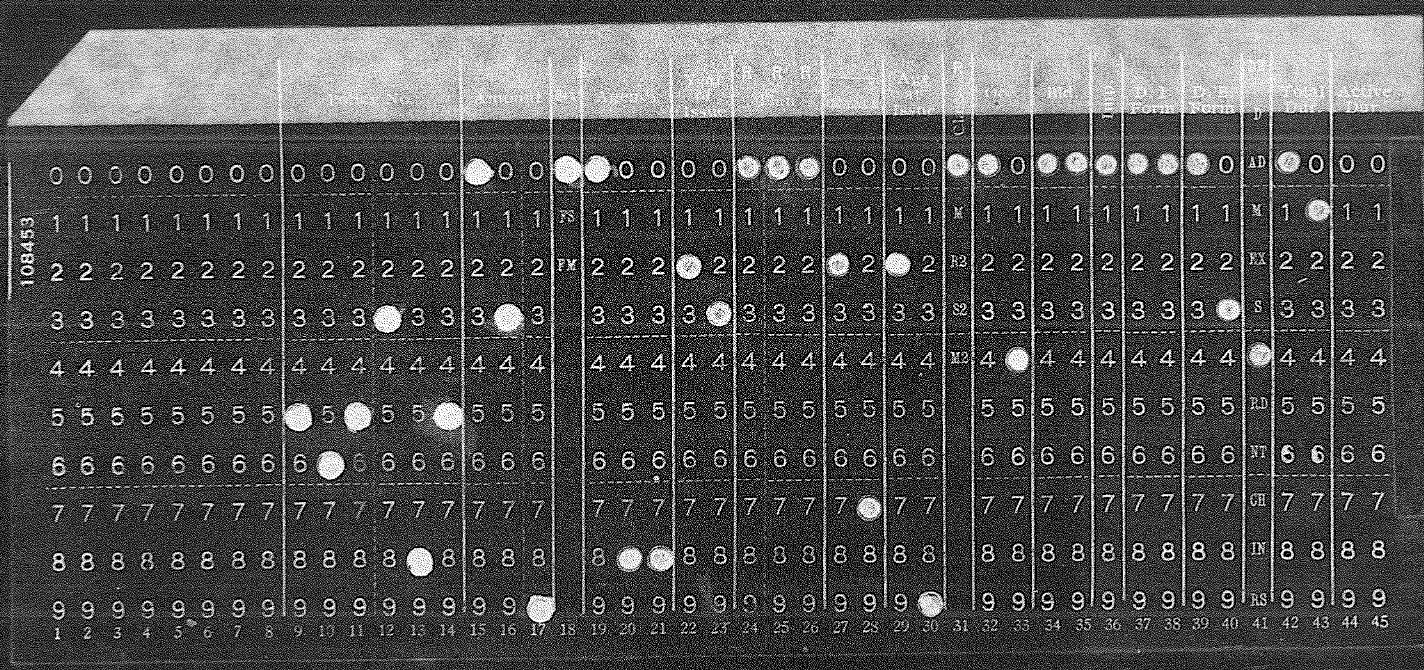

Illustration Number 2

Columns 41 to 45 are not punched unless the Disability Benefit has been terminated during the period covered by the investigation. The mode of termination of the Benefit is shown in column 41. The code used in this column in Illustration number 1, is as follows:

Disability Claim - Punch out the "R" (two spaces above the "0")
Dead (except from accident) - Punch out the "X" (one space above the "0")

Accidental Death	-	Punch	out	the	"0"
Matured	-	"	"	"	"1"
Extended Insurance		"	"	"	"2"
Surrendered	-	"	"	"	"3"
Lapsed	-	"	"	"	"4"
Reduced	-	"	"	"	"5"
Not taken	-	"	"	"	"6"
Changed	-	"	"	"	"7"

Codes 8 and 9 are not used. The mean duration in years, obtained by subtracting the calendar year of termination from the calendar year of issue, is punched in columns 42 and 43, and columns 44 and 45 are left blank, unless the mode of termination is by disability. If the policyholder becomes disabled, the current duration in policy years is punched in columns 44 and 45, and columns 42 and 43 are left blank. The current duration in years is obtained by adding one year to the completed number of policy years during which the policy has been in force on the date of disability.

The punched card shown in Illustration number 2, is for policy number 565385, for $3900, issued in 1923 in a special class by agency number 88 to an imaginary male aged 29, of normal build, and with no impairments in his family history, and on the Ordinary Life plan. The policy contained Disability Clause #3 (coded "03"), and was terminated by lapse with a mean duration of one year. All this information is indicated by the way in which the card is perforated. Only the information contained in columns 15 - 17; 22 - 23; 29 - 31; 39 - 45; is used in determining the amount exposed to risk of disability.

There are three different methods of computing this exposed to risk, the calendar year method, the policy year method, and the life year method. Under the calendar year method, each entrant into the experience is assumed to enter in the middle of the calendar year, i.e., on July 1, and the exposures are traced from the beginning to the end of each calendar year. Hence only the calendar year (not the exact date) of entry and exit are needed. The exposures are treated as exposed for the whole calendar year at their age on January 1 of the year, and the disabled as becoming claims at their age on January 1 of the year of disability.

Under the life year method, the data are traced from birthday to birthday. This method may be termed the exact method. The exact age at entry, possibly taken to one decimal place, is determined, and also the exact age at date of disability. In the case of the withdrawals and the existing, the duration, taken to the same degree of accuracy, may be determined, or an average fraction of a year of exposure may be assumed in the last year. The vast amount of work and unnecessary refinements involved in this method, prevent it from being used in practice to any extent.

Under the policy year method, the exposures are traced from the beginning to the end of each policy year, and each claim is placed in the correct policy year in which it occurs. Almost invariably, the exposures are counted from their policy anniversaries in the years in which the investigation commences, and the existing are observed to their policy anniversaries in the year in which the investigation closes. As this method is the one which is best adapted to the compilation of select tables, although the calendar year method may also be used, it is generally adopted.

The exact period to be covered by the investigation has to be determined. In the statistics to be given later, the data consisted of the experience on all policies issued with Disability Clause #3, to January 1, 1924, observed to July 1, 1926. The data were compiled by a combination of the calendar year and policy year methods. Generally speaking, under the policy year method, the observations are continued to the policy anniversaries in the last year of the investigation. When, however, the observations are made to a particular day in this last year, instead, a special adjustment has to be made, as will be pointed out later.

Compiling and Tabulating the Experience.

When the cards have all been punched, and the status of the policy, as far as the Disability Benefit is concerned, at the close of the observation period has been indicated in Columns 41 to 45, (a case in which the policy is still "existing" on this date being indicated by the fact that these columns are left blank) the cards are sorted and tabulated into certain classes. This work is automatically done by electrically operated sorting machines and tabulating machines.

In order to obtain homogeneous groups, the cards are first sorted in columns 39 and 40, the experience under each form of disability clause being separated. The statistics to be given later give the results of an experience under a particular disability clause coded 03. The experience under each clause is then sorted in column 31, and any special classes of lives which it is thought advisable to eliminate, are taken out. The cards are next sorted in column 41, the disability claims (those punched "R") being placed in one group; the existing (those not punched in this column) in another group, and all the others, which may be termed the "withdrawals", in a third group.

The "withdrawal" group is sorted according to mean duration (columns 42 and 43), and each duration is sorted according to calendar year of issue. If the investigation ends on a specified date, say July 1, 1926, which was the actual date of termination of the investigation mentioned, the policies withdrawing during a certain calendar year are kept separate. For example, those issued in 1923 and terminated with a mean duration of three years, those issued in 1922, and terminated with a mean duration of four years, etc., are separated from the others.

The assumption is made that, on the average, the policies are taken out on July 1 in their year of issue, and are also terminated in the middle of the year. This justifies the integral duration for withdrawals which is obtained by the mean duration method. For policies, however, which are terminated during the last calendar year of the observations, in the case in which the observations are taken only to a certain date in this last calendar year, this assumption does not hold. For example, if the last day of observations is July 1, 1926, those policies which are terminated by withdrawal during 1926, which will be called the "Special Withdrawals", instead of being terminated on the average on July 1, 1926, will be terminated, on the average, on April 1. This means that these policies will, on the average, be exposed to risk for an integral number of years, less three months; i. e., that they will be exposed for $2\frac{3}{4}$, $3\frac{3}{4}$, $4\frac{3}{4}$ - - - - years, instead of for 3, 4, 5, - - years, and this fact must be kept in mind in finding the exposed to risk.

After the "Withdrawals" and "Special Withdrawals" have been sorted according to duration, each duration is sorted in columns 29 and 30, according to age at issue. When so sorted the cards are passed through the electric tabulating machine. By means of a controllable system of electric circuits, which

SPECIMEN SHEET PRINTED BY THE TABULATING MACHINE

olumn(1)	Column (2)	Column (3)	Column (4)	Column (5)
19	3		10	10
20	3		10	10
21	3		10	10
24	3		25	25
26	3		50	50
28	3		20	20
29	3		90	90
30	3		50	50
31	3		20	20
32	3		20	20
33	3		20	20
35	3		100	100
42	3		150	150
44	3		25	25
46	3		100	100
40	3		50	50
47	3		300	300
50	3		60	60
53	3		50	50
18	4		10	10
49	4		10	10
20	4		25	25
24	4		40	40
25	4		10	10
26	4		110	110
27	4		10	10
28	4		25	25
39	4		25	25
30	4		20	20
32	4		45	45
34	4		100	100
35	4		100	100
37	4		50	50
38	4		20	20
39	4		100	100
42	4		130	130
43	4		10	10
46	4		10	10
18	5		30	30
20	5		20	20
24	5		20	20
27	5		70	70
32	5		25	25
33	5		25	25
35	5		70	70
42	5		100	100
45	5		100	10
48	5		20	20
51	5		50	50
52	5		10	10
373	266	1050		

ILLUSTRATION NUMBER 3

are arranged so that the machine continuously adds the amount of insurance punched on the cards, as they pass through, and automatically stops and prints the result on a suitable paper roll where either the age at issue or the mean duration changes, a list is obtained, showing for each duration and each age at issue, the amount of withdrawals, taken to the nearest hundred dollars.

Illustration number three shows a specimen sheet as printed by the tabulating machine. In the first column of this sheet the age at issue appears, in the second column, the mean duration. If the cards had been sorted properly before being tabulated, the ages at issue and durations would run smoothly, and continuously increase. In the illustration, several cases of improper sorting have been indicated. Columns four and five are duplicate columns, and should, and do, register the same, each showing the amounts of the policies withdrawing for the successive ages at issue. Column three is used to take the grand total of the figures in column four or five, and only prints once, when the cards have all been run through. The "Special Withdrawals", if any, are tabulated similarly to the regular Withdrawals, as just described.

The disability claim cards, i. e., the cards for the policyholders who have become disabled, are sorted and tabulated exactly as the withdrawals are sorted and tabulated, except that the current duration at date of disablement, which is punched on the cards, is tabulated in the second column of the tabulation sheets, and except that no group corresponding to the "Special Withdrawals" is needed.

The Existing cannot, of course, be sorted according to duration (columns 42 and 43), as they are not punched in these columns. Instead, they are sorted according to year of issue (columns 22 and 23), and age at issue (columns 29 and 30), and tabulated on the sheets in the same manner as the withdrawals, with the year of issue taking the place of the duration.

If the investigation continues to, say, July 1 in the last calendar year of observations, the existing issued in that year would, of course, have to be treated as if they were in force only three months, as they would, on the average, have been in force that length of time on July 1 if the Company's business was evenly distributed throughout the year. Any errors introduced by this assumption would likely be negligible. The policies issued in years previous to the last year of the investigation would, however, be in force, on the average, an exact number of years at the close of the investigation. In the investigation, the results of which are given below, no policies issued in the last calendar year of the investigation (1926) were included, however, as the investigation included only policies issued prior to January 1, 1924.

Obtaining the Exposed to Risk.

When the data are satisfactorily sorted and tabulated on printed rolls in accordance with the details given above, they are listed on sheets in a convenient form to obtain the Exposed to Risk.

The formula for the Exposed to Risk of Disability may be obtained as follows, in the case where the investigation is by policy years, and is terminated in the middle of a calendar year.

Let $n_{[x]}$ = total amount of insurance on policies entering at age x nearest birthday.

$\ell_{[x]+t}$ = total amount on policies entering at age x and passing out of the experience t calendar years later as existing on the closing day of the observations.

$w_{[x]+t}$ = total amount on policies entering at age x and withdrawing, previous to the last calendar year of observation, with a mean duration of t years.

$g_{[x]+t}$ = amount of insurance on those entering at age x and withdrawing, with a mean duration of t years, in the last calendar year of observations.

$i_{[x]+t}$ = amount of the disability claims on policies entering at age x and becoming claims during the $(t+1)$th policy year.

$E_{[x]+t}$ = amount of insurance on those exposed to risk of disablement during the $(t+1)$th policy year.

For each value of x, $w_{[x]}$ represents the amount of insurance on those entering at age x and withdrawing during the same calendar year. These will, on the average, be exposed to risk for one-half of the first policy year only, and will hence be equivalent to $\frac{1}{2}w_{[x]}$ exposed to risk for the full first policy year.

$$\text{Now} \quad n_{[x]} = \sum_{t=0}^{t=\infty}\left(w_{[x]+t} + g_{[x]+t} + i_{[x]+t} + e_{[x]+t}\right)$$

that is, the total amount of insurance on the policies entering the investigation at age x is equal to the sum of the amounts on policies entering at age x and classified as going out of the experience at all durations, either as existing, as withdrawals, as special withdrawals, or as disability claims. All policies must pass out of observation by one of these methods.

Let the observations terminate on July 1.
Then $E_{[x]}$ = exposed to risk for the first policy year
$$= n_{[x]} - \tfrac{1}{2}w_{[x]} - \tfrac{3}{4}e_{[x]} - \left(\tfrac{3}{4}g_{[x]} + \tfrac{1}{4}g_{[x]+1}\right)$$

since, on the basis of a uniform distribution of business, $\frac{1}{2}w_{[x]}$ and $\frac{3}{4}e_{[x]}$ and $\frac{3}{4}g_{[x]}$ may be taken as not exposed for any length of time, and the other $\frac{1}{2}w_{[x]}$ and $\frac{1}{4}e_{[x]}$ and $\frac{1}{4}g_{[x]}$ as exposed for a full year. It is also assumed that $g_{[x]+1}$ is exposed for only $\frac{3}{4}$ of a year, on the average. This is the same as assuming that $\frac{1}{4}g_{[x]+1}$ is not exposed at all, and that the remaining $\frac{3}{4}g_{[x]+1}$ is exposed to risk for a full year.

Then

$$E_{[x]+1} = E_{[x]} - \tfrac{1}{2}w_{[x]} - w_{[x]+1} - \tfrac{1}{4}e_{[x]} - e_{[x]+1} - \left(\tfrac{3}{4}g_{[x]+1} + \tfrac{1}{4}g_{[x]+2}\right) - i_{[x]}$$

$$E_{[x]+2} = E_{[x]+1} - w_{[x]+2} - e_{[x]+2} - \left(\tfrac{3}{4}g_{[x]+2} + \tfrac{1}{4}g_{[x]+3}\right) - i_{[x]+1}$$

$$E_{[x]+3} = E_{[x]+2} - w_{[x]+3} - e_{[x]+3} - \left(\tfrac{3}{4}g_{[x]+3} + \tfrac{1}{4}g_{[x]+4}\right) - i_{[x]+2}$$

$$E_{[x]+4} = E_{[x]+3} - w_{[x]+4} - e_{[x]+4} - \left(\tfrac{3}{4}g_{[x]+4} + \tfrac{1}{4}g_{[x]+5}\right) - i_{[x]+3}$$

and

$$E_{[x]+n} = E_{[x]+n-1} - w_{[x]+n} - e_{[x]+n} - \left(\tfrac{3}{4}g_{[x]+n} + \tfrac{1}{4}g_{[x]+n+1}\right) - i_{[x]+n-1}$$

or

$$E_{[x]} = n_{[x]} - \tfrac{1}{2}w_{[x]} - \tfrac{3}{4}e_{[x]} - \left(\tfrac{3}{4}g_{[x]} + \tfrac{1}{4}g_{[x]+1}\right)$$

$$E_{[x]+1} = n_{[x]} - \sum_{t=0}^{t=1}(w+e+g)_{[x]+t} - \tfrac{1}{4}g_{[x]+2} - i_{[x]}$$

$$E_{[x]+2} = n_{[x]} - \sum_{t=0}^{t=2}(w+e+g)_{[x]+t} - \tfrac{1}{4}g_{[x]+3} - \sum_{t=0}^{t=1} i_{[x]+t}$$

and

$$E_{[x]+n} = n_{[x]} - \sum_{t=0}^{t=n}(w+e+g)_{[x]+t} - \tfrac{1}{4}g_{[x]+n+1} - \sum_{t=0}^{t=n-1} i_{[x]+t}$$

The following table (Table Number 1) gives a specimen sheet for obtaining the exposed to risk of disability under an investigation conducted as described above. The figures in it are taken to the nearest hundred dollars, and are purely hypothetical.

T A B L E N U M B E R O N E.

Age at Issue _____ Total Amount Entering at this age - 981288

(1) Duration	(2) Existing Year of Issue	(2) Existing Amount	(3) Withdrawals	(4) Special Withdrawals	(5) $\frac{3}{4}$ Special Withdrawals	(6) $\frac{1}{4}$ Special Withdrawals	(7) Disability Claims	(8) Total Deduction From Prev. Year's Exposed to Risk	(9) Pol. Yr.	(9) Exposed to Risk for Policy Year Indicated
0	1926	(24478) 32637	(5533) 11066	428	321	257		30589	1	950699
1	1925	(8159) 47265	(5533) 22987	1026	769	(107) 705	3289	88814	2	861885
2	1924	43686	33406	2820	2115	768	3072	83047	3	778838
3	1923	44263	42881	3071	2303	523	2982	92952	4	685886
4	1922	40685	53600	2094	1571	466	2547	98869	5	587017
5	1921	39898	64394	1863	1397	282	2036	108007	6	479010
6	1920	39770	72063	1127	845	206	1799	114683	7	364327
7	1919	37266	60439	825	619	82	1237	99643	8	264684
8	1918	32055	58518	329	247	40	605	91465	9	173219
9	1917	26661	41695	161	121	13	404	68894	10	104325
10	1916	21037	30270	52	39	5	273	51624	11	52701
11	1915	14263	20604	20	15		250	35132	12	17569
12	1914	6737	10236				143	17116	13	453
13	1913	360	66				27	453	14	0
Total		426583	522225	13816	10362	3454	18664	981288		

In table number 1, columns (5) and (6) are so arranged that the figure in column (5) plus the figure in column (6) for the line above equals the figure in column (4). In column (8), opposite duration zero, in accordance with the formula, is put $\frac{3}{4}$ column (2) plus $\frac{1}{2}$ column (3) plus column (5) plus column (6). The figures in brackets in this line are $\frac{3}{4}$ of the existing and $\frac{1}{2}$ the withdrawals. Opposite duration one, column (8) consists of the sum of the two sets of figures in this line in column (2), (3), and (6), (the figures in brackets being $\frac{1}{4}$ the existing, $\frac{1}{2}$ the withdrawals, and $\frac{1}{4}$ the special withdrawals, respectively, each for duration zero) and the figures in columns (5) and (7). For duration two and upwards, column (8) consists of the sum of columns (2), (3), (5), (6) and (7). In each case, column (9) is equal to the exposed to risk given in column (9) in the previous line, less the amount in column (8) of the line in question.

The rate of disability for each age at issue and each policy year is found by dividing the exposed to risk of disability during that particular year into the amount of the disability claims during that year. The symbol for the rate of disability during the (t + 1)th policy year among those who entered at age x is $r_{[x]+t}$

$$\text{Hence } r_{[x]+t} = \frac{i_{[x]+t}}{E_{[x]+t}}$$

The Disability Statistics.

Table number 2 shows the form of work sheet used in obtaining the exposed to risk and the rate of disability in the investigation previously referred to several times in this thesis.

TABLE NUMBER TWO.

Age at Issue 31 — Total Amount Entering at this Age - 89717

(1) Duration	(2) Existing Year of Issue	(2) Existing Amount	(3) Withdrawals	(4) Special Withdrawals	(5) $\frac{3}{4}$Spec. Withdrawals	(6) $\frac{1}{4}$Spec. Withdrawals	(7) Disability Claims	(8) Total Deduction from prev. Year's Exposed to Risk	Policy Year	(9) Exposed to Risk of Disability	(10) Disability Claims	(11) Rate of Disability per 1000
0			(3753) 7506					3753	1	85964	235	2.7
1			(3753) 23201				235	27189	2	58775	100	1.7
2			7100			86	100	7286	3	51489	120	2.3
3	1923	17950	1945	345	259	188	120	20462	4	31027	20	.6
4	1922	16240	860	750	562	166	20	17848	5	13179		
5	1921	12680		665	499			13179	6			
Total		46870	40612	1760	1320	440	475	89717				

This investigation included entrants at each age from 15 to 55 inclusive. The exact data of this investigation for age at entry 31 are given in Table number 2. There were forty other sheets like this, one for each of the ages at entry. This table shows that in this investigation there were no existing or special withdrawals for duration zero, one, and two, (except column (6) in duration two), and no disability claims for duration zero, as these claims were sorted according to their current policy year. Columns (2), (3), (4), and (7) are summed vertically, and the grand total for all durations of the sum of these four columns is the "Total Amount Entering at this Age", given at the top of the table.

The information in this table is arranged in the same way as that in Table number 1, except that, on the right of column (8) is included a column in which the policy year is given. Then comes column (9), in which the Exposed to Risk is given, and column (10), which contains the disability claims occurring in each policy year, and column (11), which gives one thousand times the quotient of column (10) divided by column (9), i. e., the rate of disability per 1000 exposed to risk. This rate of disability is taken to one decimal place.

In Table number 3, the rates of disability per thousand, obtained from the data given on the work sheets of each of the ages at entry, are collected and shown together in convenient form. As this investigation did not, generally speaking, cover more than the first four policy years, the results shown for the fifth policy year are discarded and not used in this thesis.

The rates of disability have been compiled in the select form in order to bring out the effect of medical selection on the rate of disability. This effect is quite strong, but of short duration. By examining the ungraduated rates, it is seen that, except at ages 16, 20, 30, 31, 32, 36, 44, 48, the rate of disability is quite a lot smaller during the first policy year than it is during the second and subsequent years. It will be an approximation to the truth if we assume that the effect of medical selection wears off in one year, and that the rate of disability reaches the ultimate stage and depends only on attained age after the second policy year has been reached.

With this in mind, and also in order to have a larger experience at each age, which would eliminate some of the very large irregularities in the ungraduated rate of disability, the Exposed to Risk and Disability Claims for the second to the fourth policy years were combined, and an ungraduated rate of disability for these three policy years taken together was obtained. This stage of the work is set out in Table number 4. The resulting ungraduated rate of disability still presents irregularities, but is smoother than the ungraduated rate for the first policy year alone.

TABLE NUMBER THREE.

Unadjusted Rate of Disability per Thousand.

Age at Issue	First Policy Year	Second Policy Year	Third Policy Year	Fourth Policy Year	Fifth Policy Year
15					
16	6.6				
17		3.3	4.4		
18		1.3		1.2	8.3
19	1.0	3.7	3.4	3.5	
20	1.7	.6	.4	6.8	
21	.2	2.1	2.2	4.4	
22	.1		1.5	3.2	
23	1.6	3.7	2.3	.4	
24	1.6	3.6	1.9	6.6	4.4
25	1.2	2.8	1.4	1.8	
26	.8	2.2	2.1	1.1	4.6
27	.9	3.0	3.9	1.4	5.4
28		3.6	2.4	1.5	1.7
29	2.9	4.8	2.0	6.5	
30	4.7	2.8	1.8	2.9	
31	2.7	1.7	2.3	.6	
32	1.6	.7	1.4	.3	4.8
33	.2	4.7	4.9	3.1	3.1
34	1.2	2.2	5.2		
35	.6	1.6	3.8	3.8	3.2
36	1.0	.7	1.0	.3	
37	3.3	2.3	2.9	7.6	6.6
38	1.4	1.3	.4	1.6	3.2
39	.7	1.9	3.6	5.0	
40	2.5	1.6	6.4	7.8	1.6
41	.9	6.0	3.8	4.9	2.1
42	5.2	9.8	4.3	5.4	12.1
43	2.0	3.7	4.7	3.6	
44	4.2	4.6	2.7		2.6
45	.6	13.1	2.8	1.4	9.9
46	1.5	.7	6.2	1.3	
47	.6	2.7	11.8	7.3	
48	2.0		7.0		4.1
49	4.5	11.1	6.8	1.0	6.3
50	.5	5.0	11.3	2.3	
51	3.0		12.1	15.0	
52	5.8	10.6	14.3	10.9	
53	2.9	6.3	6.9		
54		10.3			19.0
55		10.2			

TABLE NUMBER FOUR

Total Exposed to Risk, Disability Claims, and Ungraduated Rate of Disability for the Second, Third, and Fourth Policy Years Combined.

Attained Age	Total Exposed to Risk	Total Disability Claims	Ungraduated Rate of Disability per 1000	Attained Age	Total Exposed to Risk	Total Disability Claims	Ungraduated Rate of Disability per 1000
18	9692			39	130953	210	1.6
19	25015	50	2.0	40	128722	340	2.6
20	44211	115	2.6	41	124926	295	2.4
21	60023	100	1.7	42	114754	675	5.9
22	77552	130	1.7	43	104200	740	7.1
23	90322	175	1.9	44	94747	400	4.2
24	102411	310	3.0	45	86457	420	4.9
25	110976	340	3.1	46	79730	565	7.1
26	116741	230	2.0	47	72662	70	1.0
27	123610	355	2.9	48	66073	240	3.6
28	127824	310	2.4	49	57370	280	4.9
29	137210	435	3.2	50	47192	400	8.5
30	139483	450	3.2	51	41344	180	4.4
31	139202	405	2.9	52	37590	170	4.5
32	139933	390	2.8	53	30706	270	8.8
33	140262	250	1.8	54	24136	280	11.6
34	139662	365	2.6	55	17745	160	9.0
35	144399	395	2.7	56	15502	60	
36	143102	465	3.2	57	7933		
37	143195	240	1.7	58	2950		
38	135157	295	2.2				

Graduation of a Disability Experience.

If, instead of the limited disability experience presented, we had an unlimited disability experience of the same kind, it would be natural to assume that the successive rates of disability at the successive ages would exhibit a regular progression. In other words, there would be a law observable in the progression from age to age. This assumption is not at all inconsistent with the fact that in the statistics presented great irregularities in the progression of the rates of disability from age to age occur, due to the accidental fluctuations arising from the small numbers involved in the experience.

A life insurance company's object in compiling such statistics as have been given is to obtain a guide as to probable future experience under similar conditions. It is necessary, therefore, to obtain a regular series such as would be obtained under an almost unlimited experience, in place of the irregular series actually found. This process is called "Graduation of the Series". The series before being adjusted is called the "Ungraduated Series".

Methods of Graduation.

In general, there are four methods of graduating or smoothing statistics. Under the first method, called the graphic method, a diagram is drawn to represent the ungraduated statistics. This diagram may be very irregular, or fairly smooth, according as the ungraduated statistics are quite irregular, or fairly regular. A smooth curve is then drawn by hand to represent the graduated series according to the best judgment of the graduator as to the closest fit.

The second method is called the interpolation method. Under it the series is graduated by fixing upon values at stated intervals, which may be called points of division, and interpolating for the values between the points of division by finite difference formulas. It is usual to obtain the values at the points of division very carefully, and to use osculatory interpolation for the intermediate values.

The third method is the summation method. Under this method the ungraduated data are operated upon by certain summation operators, and each term in the graduated series is obtained by combining with it, in certain proportions, the terms adjacent to it. Different summation formulas involve the use of different numbers of terms of the unadjusted data in obtaining each term of the graduated data. For example, in obtaining each graduated value in a series graduated by Spencer's 21 term formula, ten of the ungraduated terms on each side of the term in question, as well as the term itself, are used. This, of course, tends to smooth or "iron out" the irregularities in the ungraduated data.

The fourth method is the method of graduation by mathematical formulas. Under this method, a suitable curve, called a frequency distribution in some cases, is fitted to the ungraduated data. This means that it is assumed that the expression for the rate of disability can be written as a mathematical formula. In other words, if r_x denotes the true rate of disability, and x the age at the beginning of the year of disability, it is assumed that r_x may be expressed as a function of x, i. e., that $r_x = f(x)$. The problem becomes one of finding the most suitable formula to be applied to the ungraduated data. When this formula has been applied, the values given for r_x for the successive values of x are the required graduated values of the rate of disability.

Even when an ungraduated series is fairly smooth, it is difficult to graduate it by the graphic method, except for a person skilled at this task. The resulting curve depends to too large an extent on the judgment of the graduator. Owing to the very large irregularities in the ungraduated rates of disability which have been given, it is out of the question to attempt a graduation of this data by the graphic method.

Graduation by Summation.

The summation method may, however, be applied. Let E represent the operation of proceeding from one term to the next, so that $E f(x) = f(x+1)$ Also let $\delta f(x) = f(x+\frac{1}{2}) - f(x-\frac{1}{2})$, $[\overline{n}] = f(x+\frac{n}{2}) - f(x-\frac{n}{2})$, $\gamma_n f(x) = f(x+n) + f(x-n)$ and $[n] f(x) = f(x+\frac{n-1}{2}) + f(x+\frac{n-3}{2}) + \cdots + f(x-\frac{n-3}{2}) + f(x-\frac{n-1}{2})$

This indicates that $[n] f(x)$ is the sum of n terms of the ungraduated series with respect to which $f(x)$ is centrally situated. It may be proved that all these operators obey the commutative, exponential, and distributive laws to which ordinary algebraic quantities are subject. (See "Graduation of Mortality and other Tables" by Robert Henderson, page 25). Let us assume that the ungraduated series may be represented as a function of the third degree of the independent variable. This is the same as assuming that it may be written as $f(x) = k + lx + mx^2 + nx^3$

On this assumption we will have

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2} f''(x) + \frac{h^3}{6} f'''(x)$$

by Taylor's series, as fourth and higher derivatives will be zero.

$$\therefore f(x) = f(x)$$
$$f(x+1) = f(x) + f'(x) + \frac{1}{2}f''(x) + \frac{1}{6}f'''(x)$$
$$f(x+2) = f(x) + 2f'(x) + \frac{4}{2}f''(x) + \frac{8}{6}f'''(x)$$
$$\vdots$$
$$f\left(x+\frac{n-1}{2}\right) = f(x) + \frac{n-1}{2}f'(x) + \frac{(n-1)^2}{8}f''(x) + \frac{(n-1)^3}{48}f'''(x)$$
and
$$f(x-1) = f(x) - f'(x) + \frac{1}{2}f''(x) - \frac{1}{6}f'''(x)$$
$$f(x-2) = f(x) - 2f'(x) + \frac{4}{2}f''(x) - \frac{8}{6}f'''(x)$$
$$\vdots$$
$$f\left(x-\frac{n-1}{2}\right) = f(x) - \frac{n-1}{2}f'(x) + \frac{(n-1)^2}{8}f''(x) - \frac{(n-1)^3}{48}f'''(x)$$

$$\therefore f\left(x-\frac{n-1}{2}\right) + f\left(x-\frac{n-3}{2}\right) + \cdots + f\left(x+\frac{n-3}{2}\right) + f\left(x+\frac{n-1}{2}\right)$$
$$= [n]f(x) = nf(x) + \frac{\left(\frac{n-1}{2}\right)\left(\frac{n+1}{2}\right)n}{6}f''(x)$$
$$= nf(x) + \frac{n(n^2-1)}{24}f''(x)$$
and
$$\frac{[n]}{n}f(x) = f(x) + \frac{n^2-1}{24}f''(x)$$

The process of taking the average of n successive terms of the series has thus introduced an error $\frac{(n^2-1)}{24}f''(x)$ into the value of f(x).
Each time we repeat the summation process with the same number, or, it may be, a different number of terms included in the summation, an additional error of the same form is introduced. As we have assumed that $f(x)$ is a function of the third degree, $f''(x)$ is a linear function of x. In making successive summations, therefore, the error introduced by the previous summation is carried forward unchanged.

Almost all summation formulas consist of three successive summations, following a preliminary adjustment. If the first summation covers r terms, the next q terms, and the last p terms, we have the following equation:

$$\frac{[p][q][r]}{pqr}f(x) = f(x) + \frac{p^2+q^2+r^2-3}{24}f''(x) \quad \cdots\cdots\cdots (1)$$

In order that a series of the third degree may be exactly reproduced when these three summations are applied to it, a correction must be introduced.

$$\text{Now } \gamma_n f(x) = f(x+n) + f(x-n)$$

$$= f(x) + n f'(x) + \frac{n^2}{2} f''(x) + \frac{n^3}{6} f'''(x)$$
$$+ f(x) - n f'(x) + \frac{n^2}{2} f''(x) - \frac{n^3}{6} f'''(x)$$

$$= 2 f(x) + n^2 f''(x)$$

$$\therefore a\gamma_1 f(x) = 2a f(x) + a f''(x)$$

and $$b\gamma_2 f(x) = 2b f(x) + 4b f''(x)$$

and $$c\gamma_3 f(x) = 2c f(x) + 9c f''(x)$$

$$\therefore (a\gamma_1 + b\gamma_2 + c\gamma_3) f(x) = 2(a+b+c) f(x) + (a+4b+9c) f''(x)$$

or

$$(1+2a+2b+2c) f(x) - (a\gamma_1 + b\gamma_2 + c\gamma_3) f(x) = f(x) - (a+4b+9c) f''(x)$$

Putting $\left[(1+2a+2b+2c) - (a\gamma_1 + b\gamma_2 + c\gamma_3)\right] f(x)$ for $f(x)$ in equation (1), we have

$$\frac{[p][q][r]}{pqr}\left[(1+2a+2b+2c) - (a\gamma_1 + b\gamma_2 + c\gamma_3)\right] f(x)$$
$$= \frac{[p][q][r]}{pqr}\left[f(x) - (a+4b+9c) f''(x)\right]$$
$$= f(x) + \frac{p^2+q^2+r^2-3}{24} f''(x) - (a+4b+9c) f''(x)$$
$$(\text{since } f^{(iv)}(x) = 0)$$

$$= f(x) + \left\{\frac{p^2+q^2+r^2-3}{24} - (a+4b+9c)\right\} f''(x)$$

If the coefficient of $f''(x)$ is equated to zero, a series of the third degre, after a preliminary correction, upon being operated upon by $\frac{[p][q][r]}{pqr}$, will be exactly reproduced.

This condition for the third degree series to be reproduced leads to the equation

$$\frac{p^2+q^2+r^2-3}{24} = a+4b+9c \quad \cdots\cdots\cdots\cdots\cdots \quad (2)$$

Hence, if equation (2) holds,

$$\frac{[p][q][r]}{pqr}\left[(1+2a+2b+2c) - (a\gamma_1 + b\gamma_2 + c\gamma_3)\right] f(x) = f(x)$$

or

$$G(x) = \frac{[p][q][r]}{pqr}\left[(1+2a+2b+2c) - (a\gamma_1 + b\gamma_2 + c\gamma_3)\right] f(x) \quad \text{- - - - - - - -} \quad (3)$$

where G(x) is the required graduated value.

Equation (2) is indeterminate, that is, it has an indefinite number of solutions. In order that the graduated values obtained shall correspond to integral values of the variable x, it is necessary that $(p + q + r)$ should be an odd number, so that either p, q, and r must all be odd numbers, or two even and one odd. If $(p + q + r)$ was even, the graduated value of $f(x)$ would correspond either to $f(x + \frac{1}{2})$ or $f(x - \frac{1}{2})$. Each set of values of p, q, r, a, b, and c, which satisfies (2), and is in accordance with this restriction, gives a summation formula for the graduation of a series.

One of the earliest summation formulas was the one derived by Woolhouse, called Woolhouse's formula. In it, p, q, and r, are each equal to 5, and a is equal to 3, and b and c are each zero. Equation (2), hence, holds. As is seen by the general summation formula, given in equation (3), Woolhouse's formula may be written:

$$G = \frac{1}{125}[5]^3\left[1+6-3E-3E^{-1}\right]$$

$$= \frac{1}{125}[5]^3\left[1-3(E-2+E^{-1})\right] = \frac{1}{125}[5]^3\left[1-3\delta^2\right]$$

Mr. Higham suggested a modification of Woolhouse's formula. It consisted in putting $a = -1$, $b = 1$, and $c = 0$, instead of $a = 3$, $b = 0$, $c = 0$. His formula is, hence

$$G = \frac{1}{125}[5]^3[1+\gamma_1-\gamma_2] = \frac{1}{125}[5]^3([3]-\gamma_2)$$

This formula gives improved results over Woolhouse's.

Both of these formulas consist of three summations of five terms, after a preliminary adjustment. It may be proved, however, that the substitution of summations over unequal intervals for summations over equal intervals, for example, the substitution of $[4][5][6]$ for $[5]^3$, will tend to increase the smoothness of the graduated series. The same effect will be obtained by increasing the product pqr, and by extending the number of terms which the formula involves. (See Robert Henderson, "Graduation of Mortality, and Other Tables", pages 34 - 35). With this in mind, Hardy modified Higham's formula, substituting $\frac{[4][5][6]}{120}$ for $\frac{[5]^3}{125}$. This substitution, however, does not quite satisfy formula (2), as, in this case,

$$\frac{p^2+q^2+r^2-3}{24} = \frac{74}{24} = 3\tfrac{1}{12}, \text{ whereas}$$

$$a + 4b + 9c = 3$$

Other summation formulas are Karup's 19 term formula, in which $p = q = r = 5$, and $a = -\frac{3}{5}$, $b = 0$, $c = \frac{2}{5}$, giving

$$G = \frac{[5]^3}{125}\left(\frac{3}{5}+\frac{3}{5}\gamma_1-\frac{2}{5}\gamma_3\right)$$

$$= \frac{[5]^3}{625}\left(3[3]-2\gamma_3\right),$$

Spencer's 21 term formula,
in which $p = q = 5$, $r = 7$, and $a = -\frac{1}{2}$, $b = o$, $c = \frac{1}{2}$,
and which reduces to

$$G = \frac{[5]^2[7]}{350}\left(1 + [3] - \gamma_3\right), \text{ and}$$

Kenchington's 27 term formula, which is

$$G = \frac{[5][7][11]}{385}\left([3] - \gamma_3\right)$$

There are a great many summation formulas which will reproduce a curve of the third degree. The smoothness of the graduated series may be measured by the sum of the absolute values of the third differences of the graduated values, or by the sum of the squares of these values. A smoothing coefficient may be introduced, based upon the third differences, of the graduated series, which gives a value to the smoothing power of the different formulas. The formulas given above will all reproduce a series in which fourth and higher differences are zero, except Hardy's modification of Higham's formula, which introduces a second difference error. They all introduce errors in a series in which the fifth and higher, or sixth and higher, etc., differences only are zero. In general, with an increase in the smoothing power of a formula, there is combined an increase in this error. (Robert Henderson, "Graduation of Mortality and Other Statistics", page 37).

Bearing these facts in mind, it was decided to graduate both the rate of disability for the first policy year, and the rate of disability for the combined second, third, and fourth policy years, by both Higham's and Spencer's formulas. It was expected that the graduated series by Spencer's formula would be smoother than that by Higham's, but that, on the other hand, the rates of disability given by Higham's formula might be more in accordance with the actual data than that obtained by Spencer's.

The following tables, numbers 5 to 8, show the actual work of graduation and the results obtained. In the first table, number 5, which gives the rate of disability for the first policy year, graduated by Higham's formula, the ungraduated rate of disability is given in column (1), the result of operating by $[3]$ on column (1), in column (2), and the result of operating by γ_2 in column (3). Column (4) shows the difference between columns (2) and (3). Column (5) shows the result of summing Column (4) by fives, and taking one-tenth of the result. In column (6) is given the summation of column (5) in fives, and in column (7) is indicated the result of the same operation upon column (6).

Column (8), which is 2/25 of column (7), gives the graduated rates of disability. The operations in the other three tables may be similarly followed and noted.

It will be noted that one difficulty which arises in graduating a series by a summation formula is that to determine one graduated value, it is necessary to have a number of ungraduated values for a considerable interval on each side of the graduated value. There is, hence, an interval at each end of the series for which graduated values cannot be obtained by the formula. For example, in the tables just given, when Higham's formula is used, graduated rates of disability are obtained only for ages 27 to 45 inclusive, and when Spencer's formula is used, these ages are 27 to 43 inclusive, although the ungraduated values extend from age 19 to at least age 53. Various devices may be employed to extend the graduated table to cover the full range, or it may be left unextended.

One significant feature of the whole four tables is that the graduated rate of disability, both for the first year rate, and for the combined second, third, and fourth year rate, did not increase from year to year, but showed two maxima, roughly at ages 29 and 42, and a minimum at age 35, for the first year disability rate. The corresponding maxima and minimum in the disability rate for the combined experience for the second, third and fourth policy years occur at ages 30 and 43, and age 36, approximately, i. e., just one year later, as would be expected from the fact that this latter experience is arranged by attained age and excludes the first policy year only. This feature of the rate of disability was not altogether expected, but seems to be borne out by the results of other investigations such as this. (See "Report of Committee on Disability Experience", Actuarial Society of America, page 20). The decreasing rate from ages 43, for the next two ages or so, is quite probably of no significance, and is likely caused by the paucity of the data at the highest ages. It should be disregarded.

TABLE NUMBER FIVE

Graduation of Experience of First Policy Year
by
Higham's Formula.

Age	(1) Ungrad. Rate of Disability	(2) [3](1)	(3) $Y_2(1)$	(4) (2)-(3)	(5) $\frac{1}{10}$ [5](4)	(6) [5](5)	(7) [5](6)	(8) Graduated Rate $\frac{2}{25}(7)$
19	1.0							
20	1.7	2.9						
21	.2	2.0	2.6	-.6				
22	.1	1.9	3.3	-1.4				
23	1.6	3.3	1.4	1.9	.45			
24	1.6	4.4	.9	3.5	.64			
25	1.2	3.6	2.5	1.1	.54	2.04		
26	.8	2.9	1.6	1.3	.18	2.58		
27	.9	1.7	4.1	-2.4	.23	3.39	18.32	1.47
28		3.8	5.5	-1.7	.99	4.40	23.19	1.86
29	2.9	7.6	3.6	4.0	1.45	5.91	26.82	2.15
30	4.7	10.3	1.6	8.7	1.55	6.91	28.12	2.25
31	2.7	9.0	3.1	5.9	1.69	6.21	27.37	2.19
32	1.6	4.5	5.9	-1.4	1.23	4.69	24.15	1.93
33	.2	3.0	3.3	-.3	.29	3.65	19.53	1.56
34	1.2	2.0	2.6	-.6	-.07	2.69	15.97	1.28
35	.6	2.8	3.5	-.7	.51	2.29	15.01	1.20
36	1.0	4.9	2.6	2.3	.73	2.65	15.29	1.22
37	3.3	5.7	1.3	4.4	.83	3.73	17.31	1.38
38	1.4	5.4	3.5	1.9	.65	3.93	20.38	1.63
39	.7	4.6	4.2	.4	1.01	4.71	24.54	1.96
40	2.5	4.1	6.6	-2.5	.71	5.36	27.77	2.22
41	.9	8.6	2.7	5.9	1.51	6.81	31.01	2.48
42	5.2	8.1	6.7	1.4	1.48	6.96	32.40	2.59
43	2.0	11.4	1.5	9.9	2.10	7.17	32.43	2.59
44	4.2	6.8	6.7	.1	1.16	6.10	29.30	2.34
45	.6	6.3	2.6	3.7	.92	5.39	25.85	2.07
46	1.5	2.7	6.2	-3.5	.44	3.71		
47	.6	4.1	5.1	-1.0	.77	3.51		
48	2.0	7.1	2.0	5.1	.42			
49	4.5	7.0	3.6	3.4	.96			
50	.5	8.0	7.8	.2				
51	3.0	9.3	7.4	1.9				
52	5.8	11.7						
53	2.9							

T A B L E   N U M B E R   S I X

Graduation of Experience of Second, Third, and Fourth Policy Years
by
Higham's Formula.

Attained Age	(1) Ungrad. Rate of Disability	(2) [3](1)	(3) $\gamma_2(1)$	(4) (2) - (3)	(5) $\frac{1}{10}$[5](4)	(6) [5](5)	(7) [5](6)	(8) Graduated Rate $\frac{2}{25}$(7)
19	2.0							
20	2.6	6.3						
21	1.7	6.0	3.9	2.1				
22	1.7	5.3	5.6	-.3				
23	1.9	6.6	4.8	1.8	1.12			
24	3.0	8.0	3.7	4.3	1.17			
25	3.1	8.1	4.8	3.3	1.30	6.36		
26	2.0	8.0	5.4	2.6	1.45	6.64		
27	2.9	7.3	6.3	1.0	1.32	7.00	34.58	2.77
28	2.4	8.5	5.2	3.3	1.40	7.30	35.42	2.83
29	3.2	8.8	5.8	3.0	1.53	7.28	35.91	2.87
30	3.2	9.3	5.2	4.1	1.60	7.20	35.73	2.86
31	2.9	8.9	5.0	3.9	1.43	7.13	34.98	2.80
32	2.8	7.5	5.8	1.7	1.24	6.82	33.96	2.72
33	1.8	7.2	5.6	1.6	1.33	6.55	33.04	2.64
34	2.6	7.1	6.0	1.1	1.22	6.26	31.47	2.52
35	2.7	8.5	3.5	5.0	1.33	6.28	29.54	2.36
36	3.2	7.6	4.8	2.8	1.14	5.56	27.68	2.21
37	1.7	7.1	4.3	2.8	1.26	4.89	27.12	2.17
38	2.2	5.5	5.8	-.3	.61	4.69	28.52	2.28
39	1.6	6.4	4.1	2.3	.55	5.70	34.23	2.74
40	2.6	6.6	8.1	-1.5	1.13	7.68	44.56	3.56
41	2.4	10.9	8.7	2.2	2.15	11.27	57.79	4.62
42	5.9	15.4	6.8	8.6	3.24	15.22	70.09	5.61
43	7.1	17.2	7.3	9.9	4.20	17.92	79.24	6.34
44	4.2	16.2	13.0	13.2	4.50	18.00	82.83	6.63
45	4.9	16.2	8.1	8.1	3.83	16.83	80.05	6.40
46	7.1	13.0	7.8	5.2	2.23	14.86	72.39	5.79
47	1.0	11.7	9.8	1.9	2.07	12.44	65.83	5.27
48	3.6	9.5	15.6	-6.1	2.23	10.26		
49	4.9	17.0	5.4	11.6	2.08	11.44		
50	8.5	17.8	8.1	9.7	1.65			
51	4.4	17.4	13.7	3.7	3.41			
52	4.5	17.7	20.1	-2.4				
53	8.8	24.9	13.4	11.5				
54	11.6	29.4						
55	9.0							

T A B L E   N U M B E R   S E V E N.

Graduation of Experience of First Policy Year by Spencer's Formula.

Age	(1) Ungrad. Rate of Disability	(2) $\frac{1}{7}(1)$	(3) $[3](2)$	(4) $\gamma_3(2)$	(5) $(2)+(3)-(4)$	(6) $[7](5)$	(7) $\frac{1}{5}(6)$	(8) $[5](7)$	(9) Graduated Rate $\frac{1}{10}[5](8)$
17									
18			.14						
19	1.0	.14	.38						
20	1.7	.24	.41	.23	.42				
21	.2	.03	.28	.23	.08				
22	.1	.01	.28	.31	-.02				
23	1.6	.23	.47	.35	.35	2.08	.41		
24	1.6	.23	.63	.16	.70	1.13	.23		
25	1.2	.17	.51	.01	.67	1.03	.21	2.05	
26	.8	.11	.41	.64	-.12	2.20	.44	2.56	
27	.9	.13	.24	.90	-.53	3.81	.76	3.20	1.58
28			.54	.56	-.02	4.62	.92	3.81	1.81
29	2.9	.41	1.08	.34	1.15	4.33	.87	4.22	1.97
30	4.7	.67	1.47	.16	1.96	4.10	.82	4.32	2.04
31	2.7	.39	1.29	.17	1.51	4.23	.85	4.17	1.99
32	1.6	.23	.65	.50	.38	4.31	.86	3.83	1.85
33	.2	.03	.43	.81	-.35	3.87	.77	3.39	1.66
34	1.2	.17	.29	.86	-.40	2.66	.53	2.82	1.46
35	.6	.09	.40	.43	.06	1.90	.38	2.35	1.33
36	1.0	.14	.70	.13	.71	1.40	.28	2.16	1.31
37	3.3	.47	.81	.53	.75	1.94	.39	2.54	1.44
38	1.4	.20	.77	.22	.75	2.90	.58	3.20	1.67
39	.7	.10	.66	.88	-.12	4.55	.91	4.08	1.96
40	2.5	.36	.59	.76	.19	5.19	1.04	4.69	2.21
41	.9	.13	1.23	.80	.56	5.80	1.16	5.07	2.39
42	5.2	.74	1.16	.19	1.71	5.01	1.00	5.08	2.45
43	2.0	.29	1.63	.57	1.35	4.80	.96	5.01	2.40
44	4.2	.60	.98	.22	1.36	4.62	.92	4.60	
45	.6	.09	.90	1.03	-.04	4.85	.97	4.22	
46	1.5	.21	.39	.93	-.33	3.74	.75		
47	.6	.09	.59	.67	.01	3.10	.62		
48	2.0	.29	1.02	.52	.79				
49	4.5	.64	1.00	1.04	.60				
50	.5	.07	1.14	.50	.71				
51	3.0	.43	1.33						
52	5.8	.83	1.77						
53	2.9	.41							

TABLE NUMBER EIGHT

Graduation of Experience of Second, Third, and Fourth Policy Years
by
Spencer's Formula

Attained Age	(1) Ungraduated Rate of Disability	(2) $\frac{1}{7}(1)$	(3) [3](2)	(4) $]_3(2)$	(5) (2)+(3)-(4)	(6) [7](5)	(7) $\frac{1}{5}(6)$	(8) [5](7)	(9) Graduated Rate $\frac{1}{10}[5](8)$
17									
18			.29						
19	2.0	.29	.66						
20	2.6	.37	.90	.27	1.00				
21	1.7	.24	.85	.43	.66				
22	1.7	.24	.75	.73	.26				
23	1.9	.27	.94	.66	.55	5.11	1.02		
24	3.0	.43	1.14	.65	.92	4.67	.93		
25	3.1	.44	1.16	.58	1.02	4.71	.94	5.20	
26	2.0	.29	1.14	.73	.70	5.48	1.10	5.40	
27	2.9	.41	1.04	.89	.56	6.05	1.21	5.57	2.75
28	2.4	.34	1.21	.85	.70	6.10	1.22	5.66	2.78
29	3.2	.46	1.26	.69	1.03	5.50	1.10	5.63	2.78
30	3.2	.46	1.33	.67	1.12	5.17	1.03	5.53	2.76
31	2.9	.41	1.27	.71	.97	5.35	1.07	5.43	2.71
32	2.8	.40	1.07	.85	.42	5.55	1.11	5.32	2.63
33	1.8	.26	1.03	.92	.37	5.58	1.12	5.16	2.52
34	2.6	.37	1.02	.65	.74	4.97	.99	4.85	2.38
35	2.7	.39	1.22	.71	.90	4.36	.87	4.43	2.25
36	3.2	.46	1.09	.49	1.06	3.78	.76	4.05	2.17
37	1.7	.24	1.01	.74	.51	3.47	.69	4.04	2.21
38	2.2	.31	.78	.73	.36	3.71	.74	4.36	2.40
39	1.6	.23	.91	1.30	-.16	4.91	.98	5.19	2.76
40	2.6	.37	.94	1.25	.06	5.93	1.19	6.35	3.24
41	2.4	.34	1.55	.91	.98	7.95	1.59	7.70	3.75
42	5.9	.84	2.19	.93	2.10	9.25	1.85	8.76	4.20
43	7.1	1.01	2.45	1.38	2.08	10.46	2.09	9.50	4.54
44	4.2	.60	2.31	.38	2.53	10.19	2.04	9.71	4.69
45	4.9	.70	2.31	1.35	1.66	9.64	1.93	9.74	4.68
46	7.1	1.01	1.75	1.71	1.05	9.01	1.80	9.21	
47	1.0	.04	1.56	1.81	-.21	9.38	1.88	8.63	
48	3.6	.51	1.25	1.33	.43	7.79	1.56		
49	4.9	.70	2.42	1.65	1.47	7.31	1.46		
50	8.5	1.21	2.54	1.30	2.45				
51	4.4	.63	2.48	2.17	.94				
52	4.5	.64	2.53	1.99	1.18				
53	8.8	1.26	3.56						
54	11.6	1.66	4.21						
55	9.0	1.29							

Extending the Results Obtained by a Summation Formula.

An attempt was made to extend one of the tables of the graduated values of the rate of disability to complete the missing values at each end of the table. In the case of the values, at the highest ages, of the rate of disability for the first policy year, graduated by Spencer's formula, the last three values of the function which it was considered were determined with sufficient accuracy, namely the rates of disability at ages 41, 42, and 43, were taken as a basis. In accordance with the assumption made originally, it was assumed that the rate of disability for the higher ages could be expressed as a function of x (the age) of the third degree. The conditions imposed upon this function were that it should reproduce the three basic values, and also make the total expected disability claims from age 41 to age 49 (to which point only the table was extended) equal to the actual. The general expression for a function of the third degree in x such that f(a) = A, f(b) = B, f(c) = C, where A, B, & C are the given values of the function for x = a, b, and c respectively, is

$$f(x) = \frac{(x-b)(x-c)}{(a-b)(a-c)} A + \frac{(x-a)(x-c)}{(b-a)(b-c)} B + \frac{(x-a)(x-b)}{(c-a)(c-b)} C + k(x-a)(x-b)(x-c),$$

where k is a constant to be determined. (See A. Henry "Calculus and Probability" pages 26 and 27).

In the problem under discussion, $a = 41$, $b = 42$, $c = 43$, $A = 2.39$, $B = 2.45$, $C = 2.40$

$$\therefore f(x) = \frac{(x-42)(x-43)}{2}(2.39) + \frac{(x-41)(x-43)}{(-1)}(2.45) + \frac{(x-41)(x-42)}{2}(2.40) + k(x-41)(x-42)(x-43)$$

$$\therefore f(x) = 1.195(x-42)(x-43) - 2.45(x-41)(x-43) + 1.20(x-41)(x-42) + (x-41)(x-42)(x-43)k.$$

Hence the rates of disability from ages 41 to 49 inclusive become:

$$
\begin{aligned}
f(41) &= 2.39\\
f(42) &= 2.45\\
f(43) &= 2.40\\
f(44) &= 1\times2\times3k + 2.39 - 7.35 + 7.20 = 2.24 + 6k\\
f(45) &= 2\times3\times4k + 7.17 - 19.60 + 14.40 = 1.97 + 24k\\
f(46) &= 3\times4\times5k + 14.34 - 36.75 + 24.00 = 1.59 + 60k\\
f(47) &= 4\times5\times6k + 23.90 - 58.80 + 36.00 = 1.10 + 120k\\
f(48) &= 5\times6\times7k + 35.85 - 85.75 + 50.4 = .50 + 210k\\
f(49) &= 6\times7\times8k + 50.19 - 117.60 + 67.2 = -.21 + 336k
\end{aligned}
$$

To obtain the expected disability claims for ages 41 to 49, the above expressions for the rate of disability are multiplied by the exposed to risk of disability for the first policy year. The work is shown in the following table (Table number 9).

TABLE NUMBER NINE.

Experience of First Policy Year

Age (1)	Exposed to Risk (2)	Actual Disability Claims (3)	Expected Disability Claims (4)
41	54265	50	130
42	51490	270	126
43	49838	100	120
44	40810	170	91 + 244.86k
45	41144	25	81 + 987.456k
46	33793	50	54 + 2027.580k
47	31505	20	35 + 3780.60k
48	24975	50	12 + 5244.75k
49	22038	100	- 5 + 7404.768k
Total		835	644 + 19690.014k

Column (4) of the above table gives the expected disability claims, involving the undetermined constant k, obtained by multiplying the values of f(41), f(42), f(43), f(44) . . . , f(49), by the exposed to risk for the same age indicated in column (1). The condition to determine k is that the total expected claims should be equal to the total actual claims over the period. This condition leads to the equation

$$644 + 19690.014k = 835$$

or

$$19690.014k = 191$$

$$\therefore \quad k = .009700348$$

Substituting this value of k in the expressions for the rates of disability (f(x)) in terms of k, we obtain the graduated rates per thousand shown in column (4) of Table number 10.

This table checks the work of extending Spencer's formula, by showing that these values of the rate of disability, when applied to the exposed to risk for the first policy year, for ages 41 to 49 inclusive, give total expected claims of 835, exactly reproducing the total actual claims for these ages.

TABLE NUMBER TEN.

Experience of First Policy Year

Age (1)	Exposed to Risk (2)	Actual Disability Claims (3)	Graduated Rate of Disability (4)	Expected Disability Claims (5)
41	54265	50	2.39	130
42	51490	270	2.45	126
43	49838	100	2.40	120
44	40810	170	2.30	94
45	41144	25	2.20	91
46	33793	50	2.17	73
47	31505	20	2.26	71
48	24975	50	2.54	63
49	22038	100	3.05	67
Total		835		835

By the same method, the tables graduated by both Higham's and Spencer's formulas, could be extended to cover the lower and higher ages.

Graduation by Mathematical Formulas.

The fourth method of graduation, mentioned on page 18, is by a mathematical formula. In order to apply this method of graduation, the unadjusted data was grouped together according to decennial groups of ages at issue, the groups chosen being ages 15 - 24; 25 - 34; 35 - 44; 45 and over.

The following were the results:

TABLE NUMBER ELEVEN.

Ages at Issue 15 - 24 Inclusive.

Policy Year	Exposed to Risk of Disability	Disabled	Rate of Disability per 1000
1	416206	435	1.0
2	262171	555	2.1
3	229589	405	1.8
4	137481	525	3.8
5	57601	80	1.4

Ages at Issue 25 - 34 Inclusive

1	848196	1395	1.6
2	573072	1625	2.8
3	498862	1110	2.2
4	305686	855	2.8
5	132875	260	2.0

Ages at Issue 35 - 44 Inclusive

1	630873	1265	2.0
2	472985	1415	3.0
3	418597	1365	3.3
4	263207	1075	4.1
5	118500	365	3.1

Ages at Issue 45 and over

1	227409	405	1.8
2	178331	1035	5.8
3	158716	1160	7.3
4	98745	335	3.4
5	41503	170	4.1

Two tables were then formed, one giving the exposed to risk, disability claims, and rate of disability per 1000 for the experience during the first policy year, for each of the above four groups, and the other giving the corresponding data for the aggregate of the second, third, and fourth policy years (divided into the four above groups) according to age at issue, and not attained age. The experience of the fifth policy year was discarded. This information is given in Table number 12.

TABLE NUMBER TWELVE.

First Policy Year.

Ages at Issue	Equivalent Age	Exposed to Risk	Disabled	Rate of Disability per 1000
15 - 24	21.4	416206	435	1.0
25 - 34	29.5	848196	1395	1.6
35 - 44	38.9	630873	1265	2.0
45 & over	48.4	227409	405	1.8
Total		2122684	3500	1.6

Aggregate of Second, Third, and Fourth Policy Years.

Ages at Issue	Equivalent Attained Ages	Exposed to Risk	Disabled	Rate of Disability per 1000
15 - 24	23.1	629241	1485	2.4
25 - 34	31.3	1377620	3590	2.6
35 - 44	40.8	1154789	3855	3.3
45 & over	50.5	435792	2530	5.8
Total		3597442	11460	3.2

For each of the ten year groups in the above table, an equivalent attained age had to be determined, such that the assumption could be made without sensible error that the experience, instead of being distributed over ten or even twelve years of attained age, was all concentrated at this equivalent age.

For the table giving the experience of the first policy year, this equivalent attained age was obtained by taking the weighted arithmetic mean of the ages entering into each group, the weights used being the amount of the exposed to risk at each of the ages at issue. Thus, to obtain the equivalent age for the 15 - 24 group, the exposed to risk at age 15 was multiplied by 15, the exposed to risk at age 16 was multiplied by 16, and so on. The sum of these ten items was taken, and this sum was divided by the total exposed to risk for that age group. The results are shown in Table number twelve. The same method was followed in the table showing the experience of the second, third, and fourth policy years, great care being taken to see that the component parts of the exposed to risk were multiplied by their correct attained age.

The experiences for the two sets of policy years were then graduated separately. It was assumed that the rate of disability in each of these tables, at each age, could be expressed by Makeham's second modification of Gompertz's law, i. e., it was assumed that $r_x = A + Hx + Bc^x$

where r_x is the graduated rate of disability at age x. This is the same as assuming that the rate of disability consists of three elements, one a constant, another increasing in arithmetical proportion throughout life, and the third increasing in geometrical proportion throughout life.

Determining the Constants in Makeham's Formula.

The above formula contains four constants, A, H, B, c. As there were four rates of disability at four different equivalent ages given in the experience for each set of policy years, it is seen that the constants in the formula could have been determined so as to exactly reproduce the rates for the four equivalent ages, in each case.

The Makeham curve, $r_x = A + Hx + Bc^x$, was first fitted to the disability experience of the first policy year. An attempt was made to determine the values of the four constants by means of four equations obtained by substituting the values of r_x and x, for each of the four attained ages, in the equation of the curve. It was almost impossible to solve the resulting four equations, due to the fact that the attained ages were not equidistant.

Four equidistant ages were, therefore, determined upon. The two central equivalent ages were kept unchanged, and the equal age intervals were taken as 9.4 years. Hence the four equidistant ages chosen were 20.1, 29.5, 38.9, 48.3. A rough interpolation of the rate of disability, to determine its approximate value at these four ages, was made. It was assumed that the unadjusted rate of disability at age 20.1 was .9, and that at age 48.3 was 1.8. The rates at ages 29.5 and 38.9 were already given as 1.6, and 2.0 per thousand, respectively.

Hence the following equations were obtained:

$$A + 20.1H + Bc^{20.1} = .9 \qquad (1)$$

$$A + 29.5H + Bc^{29.5} = 1.6 \qquad (2)$$

$$A + 38.9H + Bc^{38.9} = 2.0 \qquad (3)$$

$$A + 48.3H + Bc^{48.3} = 1.8 \qquad (4)$$

leading to

$$(1)-(2) \qquad -9.4H + Bc^{20.1}(1-c^{9.4}) = -.7 \qquad (5)$$

$$(2)-(3) \qquad -9.4H + Bc^{29.5}(1-c^{9.4}) = -.4 \qquad (6)$$

$$(3)-(4) \qquad -9.4H + Bc^{38.9}(1-c^{9.4}) = .2 \qquad (7)$$

$$(5)-(6) \qquad Bc^{20.1}(1-c^{9.4})^2 = -.3 \qquad (8)$$

$$(6)-(7) \qquad Bc^{29.5}(1-c^{9.4})^2 = -.6 \qquad (9)$$

$$c^{9.4} = 2$$

$$\therefore 9.4 \log_{10} c = \log_{10} 2 = .30103$$

$$\therefore \log c = .03202$$

$$\therefore c = 1.0765$$

$$\text{Now } 20.1 \log c = .643602$$

$$\therefore c^{20.1} = 4.4015$$

Therefore, from (8), $4.4015\,B = -.3$

$$\therefore B = -.0681586$$

Now, from (5), and (8), $Bc^{20.1}(1-c^{9.4}) = .3$

and $-9.4H = -1.0 \qquad \therefore H = .10638$

Now from (1), $A + 2.138238 - .3 = .9 \qquad \therefore A = -.938238$

Hence the formula for the rate of disability at age x becomes:

$$r_x = .10638x - .938238 - (.0681586)(1.0765)^x$$

To test the formula, let us find $r_{21.4}$ Now

$r_{21.4} = 2.2765 - .9382 - (.068159)(4.8444) = 1.0081,$

whereas the value of $r_{21.4}$ should be 1.0.

Modifying Makeham's Formula.

It was seen that this formula was unsatisfactory, becaues, by it, the rate of disability from age 41 onward continues to decrease, and at age 55 becomes only .981 per thousand, which is, of course, rather unbelievable. The reason for this peculiarity is that the ungraduated average rate of disability for the age group 45 and over is actually less than that for the age group 35 - 44. As, however, the experience of the second, third, and fourth years shows that the rate for the former age group should be higher than that for the latter, it was decided to add a term to the formula, of the form $a^{(x-M)^n}$,

where "a" is a constant just greater than one, "M" is an age near the extreme age limit of the table, say close to 50, and "n" is an integer less than 10. It may be seen that for such values of a, M and n, the expression $a^{(x-M)^n}$ will always be a positive quantity, less than unity for ages less than M, unity at age M, and increasing very rapidly thereafter. This term will be small for ages much less than M, and will increase very gradually until an age just less than M is reached. This is the sort of a function required, as its effect should be negligible except at, say, the last ten ages, for which the rates of disability are to be obtained.

As stated in the disability clause quoted at the beginning of this thesis, the insured has to become disabled before age sixty for the disability benefit to apply. This means, in practice, that the disability benefits are allowed only on policies taken out at ages less than 56, so that for the experience of the first policy year, the highest age is 55. This restriction holds in the case of the experience presented, and hence the limiting age is age 55.

An expression had to be obtained which would give the desired results for values of x up to 55. No notice was taken of what the results would be for higher values. At age "M", the value of the term $a^{(x-M)^n}$ is unity. After much experiment, age "M" was taken as 53, "a" as 1.0765 (to agree with "c"), and "n" as 3, the additional term, hence, becoming $(1.0765)^{(x-53)^3}$, and the complete formula for the rate of disability becoming

$$r_x = .10638x - .938238 - (.0681586)(1.0765)^x + (1.0765)^{(x-53)^3}$$

After the formula had been obtained, it was tested at a few ages to check its general results. The ages chosen were 20, 30, 40, 45, 50, and 55. The results were as follows:

$$r_{20} = 2.1276 - .9382 - .2978 + .0007$$
$$= .8923$$

$$r_{30} = 3.1914 - .9382 - .6225 + .0062$$
$$= 1.6369$$

$$r_{40} = 4.2552 - .9382 - 1.3011 + .0564$$
$$= 2.0723$$

$$r_{45} = 4.7871 - .9382 - 1.8811 + .1704$$
$$= 2.1382$$

$$r_{50} = 5.3190 - .9382 - 2.7197 + .5150$$
$$= 2.1761$$

$$r_{55} = 5.8509 - .9382 - 3.9321 + 1.5564$$
$$= 2.5370$$

As was to be expected, these graduated rates are very slightly in excess of the ungraduated, but nevertheless agree quite well for ages up to 40, or even higher. It will be noted that, owing to the nature of the formula used to express the rate of disability, this rate remains practically stationary from age 42 to age 49, and increases at a moderate rate only for the last two or three ages in the table.

The graduated rates of disability for the first policy year, obtained by this formula, for all ages from 15 to 55 inclusive, are given in Table number thirteen.

TABLE NUMBER THIRTEEN

Rate of Disability per 1000
Experience of First Policy Year Only
Graduated by Makeham's Formula, Modified.

Age	Rate of Disability per 1000	Age	Rate of Disability per 1000	Age	Rate of Disability per 1000	Age	Rate of Disability per 1000
15	.4517	25	1.2930	35	1.9038	45	2.1382
16	.5425	26	1.3667	36	1.9459	46	2.1428
17	.6319	27	1.4383	37	1.9839	47	2.1469
18	.7200	28	1.5073	38	2.0177	48	2.1521
19	.8069	29	1.5735	39	2.0471	49	2.1608
20	.8923	30	1.6369	40	2.0723	50	2.1761
21	.9760	31	1.6972	41	2.0931	51	2.2019
22	1.0581	32	1.7542	42	2.1063	52	2.2434
23	1.1383	33	1.8077	43	2.1224	53	2.3085
24	1.2166	34	1.8577	44	2.1317	54	2.4013
						55	2.5370

Experience of the Second, Third, and Fourth Policy Years.

The experience of the second, third, and fourth policy years combined was also graduated by Makeham's second modification of Gompertz's formula in a similar manner. Bearing in mind the fact that in order to obtain the constant "c", four equidistant equivalent attained ages had to be taken, the same difference between these attained ages of 9.4 years was assumed, and the attained ages used were 22.0, 31.4, 40.8, and 50.2 years. The actual ungraduated rates of disability are given in table number twelve for attained ages 23.1, 31.3, 40.8, and 50.5. By a rough interpolation, the ungraduated rate at age 22.0 was taken as 2.3, that at 31.4 as 2.6, and that at age 50.2 as 5.7 per thousand.

The formula for r_x, the graduated rate of disability per thousand, was taken as before, as

$r_x = A + Hx + Bc^x$, where x is the attained age of the insured at the beginning of the policy year during which disability occurred. By solving the four equations for the rate of disability at ages 22.0, 31.4, 40.8, and 50.2, for the four constants, A, H, B, and c, the following values of the constants were obtained:

$$
\begin{aligned}
A &= 2.00389 \\
H &= .011994 \\
B &= .002933 \\
c &= 1.14893 \\
\text{and log } B &= \bar{3}.4672901 \\
\text{and log } c &= .0603080
\end{aligned}
$$

Substituting these values in the equation for the rate of disability, it becomes

$$r_x = 2.0039 + .01199x + .0029(1.1489)^x$$

The rates produced by this formula agree quite well with the ungraduated rates at the four attained ages, the graduated rates at the nearest integral ages being

$$
\begin{aligned}
r_{23.0} &= 2.3512 \\
r_{31.0} &= 2.5928 \\
r_{41.0} &= 3.3663 \\
r_{50.0} &= 5.6421
\end{aligned}
$$

The exposed to risk for the aggregate experience of the second, third, and fourth policy years extended from attained ages 18 to 58 inclusive. The graduated rates for these ages, obtained from the above formula, are given in table number fourteen.

TABLE NUMBER FOURTEEN

Experience of Second, Third, and Fourth Policy Years
Graduation by Makeham's Formula.

Age	Graduated Rate of Disability per 1000	Age	Graduated Rate of Disability per 1000	Age	Graduated Rate of Disability per 1000	Age	Graduated Rate of Disability per 1000
18	2.2554						
19	2.2728	29	2.5161	39	3.1311	49	5.2361
20	2.2908	30	2.5526	40	3.2414	50	5.6421
21	2.3099	31	2.5928	41	3.3663	51	6.1068
22	2.3299	32	2.6371	42	3.5080	52	6.6389
23	2.3512	33	2.6862	43	3.6691	53	7.2485
24	2.3739	34	2.7410	44	3.8523	54	7.9471
25	2.3981	35	2.8021	45	4.0611	55	8.7480
26	2.4241	36	2.8704	46	4.2990	56	9.6663
27	2.4522	37	2.9472	47	4.5708	57	10.7197
28	2.4828	38	3.0336	48	4.8812	58	11.9260

Criteria of a Good Graduation

The four principal methods of graduating a series of statistics have now been discussed somewhat, and we have obtained three graduations of the disability experience presented in this thesis for the experience of the first policy year, and for the combined experience of the second, third, and fourth policy years.

After a graduation has been made of a set of statistics, it is usually tested with respect to smoothness and to closeness of agreement with the ungraduated data. The accepted criterion of the smoothness of a graduated table is the extent of the third differences of the graduated series. In order to have a means of comparing the smoothness of different graduations, the sum of the squares of the third differences, or the actual sum of these differences without regard to sign, taken over the whole series, is taken as a numerical indication of the smoothness of the graduation. In the case of the graduations presented in this thesis, the latter criterion was adopted.

The graduation by Higham's summation formula of the rate of disability for the first policy year covered ages 27 to 45 inclusive; the graduation by Spencer's formula (including the extension) covered ages 27 to 49; and the graduation by the Makeham formula covered ages 15 to 55 inclusive. The third differences of these graduated rates were, hence, extracted for the common ages of 27 to 45 and their totals, without regard to sign, obtained. The same procedure was followed for the graduation of the second, third, and fourth policy years' experience, and the differences were taken over the same common period of ages from 27 to 45 inclusive. The results were as follows:

Sum of Third Differences:

	First Policy Year	Second, Third & Fourth Policy Years
Higham's Formula	1.54	1.40
Spencer's Formula	.73	.61
Makeham's Formula	.21	.14

These results are fairly consistent for the three formulas as between the first year's experience, and the later experience. A Makeham graduation produces by far the smoothest series. The third differences by Spencer's Formula average $3\frac{1}{2}$ or 4 times as great, and those by Higham's Formula, from $7\frac{1}{2}$ to 10 times as great. In fact, it is evident that when a formula of the same type as Makeham's law is used, the third differences represent those arising from the law itself together with those resulting from dropped fractions. If the ungraduated series was graduated by the graphic method, the smoothness of the resulting series could be made as great as the graduator wished.

With respect to the closeness of fit of the graduated data with the ungraduated, the requirements usually made are that the total number and the first and second moments about any assigned point shall be approximately the same in the graduated series as in the ungraduated series, and that the departures between the graduated and ungraduated values, both for individual ages and for groups of ages, shall not greatly exceed those to be expected according to the theory of probability.

In order to make a comparison between the different graduations table number fifteen and sixteen were drawn up, the first for the first year's experience, and the second for the combined second, third, and fourth years' experience. In forming each of these tables, the exposed to risk at each age was multiplied by the graduated rate of disability per thousand, and the expected number of disability claims thus found. The difference between these expected number of claims and the actual number of claims, positive if the expected exceeded the actual, was recorded under the head of deviation. This column of deviations was then summed from the top down to obtain the accumulated deviation from the beginning of the table to each successive age. The smallness of the numbers in this column of accumulated deviations, together with the frequency of the changes in sign of both the individual deviations and the accumulated deviations, and the extent to which positive and negative deviations balanced one another, were taken as indications of the extent of the agreement between the total number of expected claims and their first and second moments and the corresponding figures with regard to the actual claims.

Table number 15 compares the actual and the expected claims for the first policy year for the three formulas. For Higham's formula, ages 27 to 45 are indicated, for Spencer's, 27 to 49, and for Makeham's (modified), 15 to 55. The deviation in the total claims is smallest by Higham's formula, being -3, slightly larger by Spencer's formula, (-29), and larger still by Makeham's (modified), being (-44). All of these deviations, however, are quite small, and may be taken as being within the limits of error, and the three formulas may be taken as reproducing the actual claims, in total, sufficiently accurately. Under Higham's formula, the accumulated deviations for eleven ages are positive, and for only eight ages are negative, showing that, generally speaking, the total accumulated expected claims are too high over more than one half of the table. However, the fact that the sign of the accumulated deviation changes seven times, shows that the error in this respect cannot be large. The individual deviations are positive at twelve ages and negative at seven ages. In the graduation by Spencer's formula, these peculiarities are slightly noticeable, there being twelve ages

T A B L E   N U M B E R   F I F T E E N.

Comparison of Actual to Expected - Experience of the First Policy Year.

		Graduation by Higham's Formula			Graduation by Spencer's Formula			Graduation by Makeham's Formula		
Age	Actual Claims	Expected Claims	Deviation (Exp. Act.)	Accumulated Deviation	Expected Claims	Deviation	Accumulated Deviation	Expected Claims	Deviation	Accumulated Deviation
15	0							0	0	0
16	50							4	-46	-46
17	0							8	+8	-38
18	0							17	+17	-21
19	35							29	-6	-27
20	85							44	-41	-68
21	10							60	+50	-18
22	10							75	+65	+47
23	120							87	-33	+14
24	125							94	-31	-17
25	100							106	+6	-11
26	65							114	+49	+38
27	75	121	+46	+46	130	+55	+55	119	+44	+82
28	0	165	+165	+211	160	+160	+215	134	+134	+216
29	260	190	-70	+141	174	-86	+129	139	-121	+95
30	410	196	-214	-73	178	-232	-103	143	-267	-172
31	235	188	-47	-120	171	-64	-167	146	-89	-261
32	130	158	+28	-92	151	+21	-146	143	+13	-248
33	20	131	+111	+19	140	+120	-26	152	+132	-116
34	100	106	+6	+25	121	+21	-5	154	+54	-62
35	50	96	+46	+71	107	+57	+52	153	+103	+41
36	80	97	+17	+88	104	+24	+76	154	+74	+115
37	235	98	-137	-49	103	-132	-56	141	-94	+21
38	100	113	+13	-36	115	+15	-41	139	+39	+60
39	50	136	+86	+50	136	+86	+45	142	+92	+152
40	160	145	-15	+35	144	-16	+29	135	-25	+127
41	50	135	+85	+120	130	+80	+109	114	+64	+191
42	270	133	-137	-17	126	-144	-35	108	-162	+29
43	100	129	+29	+12	120	+20	-15	106	+6	+35
44	170	95	-75	-63	94	-76	-91	87	-83	-48
45	25	85	+60	-3	91	+66	-25	88	+63	+15
46	50				73	+23	-2	72	+22	+37
47	20				71	+51	+49	68	+48	+85
48	50				63	+13	+62	54	+4	+89
49	100				67	-33	+29	47	-53	+36
50	10							42	+32	+68
51	50							37	-13	+55
52	70							27	-43	+12
53	30							24	-6	+6
54	0							19	+19	+25
55	0							19	+19	+44
Total	3500	2517	-3	+365	2769	+29	+138	3544	+44	+582

TABLE NUMBER SIXTEEN.

Comparison of Actual to Expected Claims
Experience of the Second, Third, and Fourth Policy Years.

		Graduation by Higham's Formula			Graduation by Spencer's Formula			Graduation by Makeham's Formula		
Attained Age	Actual Claims	Expected Claims	Deviation (Exp. Act.)	Accumulated Deviation	Expected Claims	Deviation	Accumulated Deviation	Expected Claims	Deviation	Accumulated Deviation
18	0							22	+22	+22
19	50							57	+7	+29
20	115							101	-14	+15
21	100							139	+39	+54
22	130							181	+51	+105
23	175							212	+37	+142
24	310							243	-67	+75
25	340							266	-74	+1
26	230							283	+53	+54
27	355	342	-13	-13	340	-15	-15	303	-52	+2
28	310	362	+52	+39	355	+45	+30	317	+7	+9
29	435	394	-41	-2	381	-54	-24	345	-90	-81
30	450	399	-51	-53	385	-65	-89	356	-94	-175
31	405	390	-15	-68	377	-28	-117	361	-44	-219
32	390	381	-9	-77	368	-22	-139	369	-21	-240
33	250	370	+120	+43	353	+103	-36	377	+127	-113
34	365	352	-13	+30	332	-33	-69	383	+18	-95
35	395	341	-54	-24	325	-70	-139	405	+10	-85
36	465	316	-149	-173	311	-154	-293	411	-54	-139
37	240	311	+71	-102	316	+76	-217	422	+182	+43
38	295	308	+13	-89	324	+29	-188	410	+115	+158
39	210	359	+149	+60	361	+151	-37	410	+200	+358
40	340	456	+116	+176	417	+77	+40	417	+77	+435
41	295	577	+282	+458	468	+173	+213	421	+126	+561
42	675	644	-31	+427	461	-214	-1	403	-272	+289
43	740	661	-79	+348	473	-267	-268	382	-358	-69
44	400	628	+228	+576	444	+44	-224	365	-35	-104
45	420	553	+133	+709	405	-15	-239	351	-69	-173
46	565	462	-103	+606				343	-222	-395
47	70	383	+313	+919				332	+262	-123
48	240							323	+83	-50
49	280							300	+20	-30
50	400							266	-134	-164
51	180							252	+72	-92
52	170							250	+80	-12
53	270							223	-47	-59
54	280							192	-88	-147
55	160							155	-5	-152
56	60							150	+90	-62
57	0							85	+85	+23
58	0							35	+35	+58
Total	11560	8989	+919	+3790	7196	-239	-1812	11618	+58	-446

at which the accumulated deviation is positive, and eleven at which it is negative, and only six changes in sign. The individual deviations are positive at fifteen ages, and negative at only eight ages. As the total deviation is only (+29), the average of the individual negative deviations must be greater than that of the positive deviations. In the graduation by Makeham's formula (modified), the number of ages at which the accumulated deviation is positive is twenty-six, and the number at which it is negative is fourteen. There are seven changes in sign in the accumulated deviation. At twenty-four ages, the individual deviations are positive, and at only sixteen are they negative. The same characteristics as were seen in the other two graduations, therefore, occur in the Makeham (modified) graduation, and to about the same extent.

The table for the experience of the combined second, third, and fourth policy years (table number sixteen) shows a less satisfactory result. In the graduation by Higham's formula, covering the ages 27 to 47 inclusive, the expected number of claims exceeds the actual by 919. At age 39, this total excess of expected over actual is only 60, but for every age from there up, the ungraduated rate is lower than the graduated, so that the accumulated deviation (being positive) continually increases. The opposite is true in the graduation by Spencer's formula. Under this formula, taking into account the whole range of ages from 27 to 45, the expected claims are 239 less than the actual. At age 42, the accumulated deviation is (-1), but for the last three ages the expected claims are 238 too few. The result under the Makeham graduation, however, is quite satisfactory. For the full range of ages of 18 to 58, the total expected claims are only 58 in excess of the actual. As stated, Higham's formula, and Spencer's to a lesser degree, give results quite unsatisfactory in this regard.

Although Higham's formula is so disappointing along this line, its results compare well with those of Spencer's formula, in other respects. Under Higham's formula, at twelve ages the accumulated deviations are positive, and at nine, negative. There are five changes in sign in these accumulated deviations. There are ten positive individual deviations, and eleven negative ones, but the average individual positive deviation is much larger than the average negative deviation. Under Spencer's formula, there are eight positive individual deviations, and eleven negative ones, and four changes in sign in the accumulated deviation. For sixteen of the nineteen ages, however, the accumulated deviation is negative. Makeham's formula gives eleven positive accumulated deviations in succession from age 18 on, then eight consecutive negative ones, and then six positive ones, and fourteen negative ones, and two positive ones for the last two ages of the table. This gives a total of nineteen positive accumulated deviations, and twenty-two negative ones. As will be seen, there are only four changes in sign. At twenty-three ages, the individual deviations are positive, and at eighteen only are they negative. The fact that the accumulated deviations continue either positive or negative for such long periods at a time indicates that the Makeham curve fitted to the data does not at all closely follow the original curve representing the ungraduated data. The accumulated deviation at one point is (-561).

Comparison of the Three Graduations of the Data.

In order to be able to make a more definite comparison between the results obtained by the three graduations given, tables number seventeen and eighteen were compiled, showing the results taken over the common period of eighteen years only, from age 27 to age 45, in each case.

TABLE NUMBER SEVENTEEN

Experience of the First Policy Year

Graduation By	(1)	(2)	(3)	(4)	(5)
Higham's	-3	+365	1387	1271	7
Spencer's	-25	+0	1475	1420	5
Makeham's (Modified)	-23	-604	1659	1924	4

TABLE NUMBER EIGHTEEN

Experience of the Second, Third, and Fourth Policy Years.

Graduation By	(1)	(2)	(3)	(4)	(5)
Higham's	+709	+2265	1619	3467	5
Spencer's	-239	-1812	1635	2378	4
Makeham's	-227	-664	1951	3726	2

In each of the above tables, column (1) shows the sum of the deviations with regard to sign, (for ages 27 to 45), or the deviation in the total of the disability claims. Column (2) in each case shows the sum of the accumulated deviations with regard to sign; column (3) shows the sum of the individual deviations without regard to sign; column (4), the sum of the accumulated deviations without regard to sign; column (5), the number of changes in the sign of the accumulated deviations.

These tables show that the total accumulated deviation is satisfactory for all three graduations, for the first policy year's experience, but unsatisfactory for the later experience. The Makeham graduation gives the best all-around results in this connection, with the Spencer graduation a rather close second. The second column gives a relative measure of the extent to which the second moment of the graduated data differs from that of the ungraduated. Possibly there is not much to choose between the Makeham and the Spencer graduation in this respect. Columns (3) and (4) show that, as far as agreement with the original facts for individual ages is concerned, Spencer's graduation is best, Higham's next, and Makeham's worst. The reason for this is that Makeham's formula fails to show the decrease in the rate of disability from ages 30 to 35, shown by the other two formulas. This characteristic, which does not seem to be accidental, might have been brought out had different decennial groups, such as ages 20 - 29; 30 - 39; 40 - 49; etc., been chosen in determining the values of the constants in Makeham's formula. Column (5) gives the advantage to Higham's formula, over the other two.

Taking everything into account, the Makeham graduation can be considered most satisfactory only if it is agreed that the decrease in the rate of disability from ages 30 to 35 is purely accidental, and should not be reproduced in the graduated series. If this decrease is significant, Spencer's formula appears to give the best graduation, It should not, however, be accepted without modification at certain ages, as it appears fairly certain that the decreasing rates of disability at the two or three highest ages are purely accidental, and caused by scanty data.

Possibly in this, as in other cases where the data are not very extensive, and where the ungraduated results are very irregular, a combination of two or more of the four methods of graduation will produce the best results, the series being divided into two or more sections, if necessary, and different formulas being applied to the different sections.

Table number nineteen presents the graduated rates of disability obtained by the three formulas, showing how the first year rate runs into the ultimate rate of the second, third, and fourth

T A B L E N U M B E R N I N E T E E N

Graduated Rates of Disability per Thousand.

Age at Entry	By Higham's Formula: For First Policy Year	By Higham's Formula: For Second and Subsequent Years	By Spencer's Formula: For First Policy Year	By Spencer's Formula: For Second and Subsequent Years	By Makeham's Formula (modified): For First Policy Year	By Makeham's Formula (modified): For Second and Subsequent Years
15					.45	
16					.54	
17					.63	2.26
18					.72	2.27
19					.81	2.29
20					.89	2.31
21					.98	2.33
22					1.06	2.35
23					1.14	2.37
24					1.22	2.40
25					1.29	2.42
26		2.77		2.75	1.37	2.45
27	1.47	2.83	1.58	2.78	1.44	2.48
28	1.86	2.87	1.81	2.78	1.51	2.52
29	2.15	2.86	1.97	2.76	1.57	2.55
30	2.25	2.80	2.04	2.71	1.64	2.59
31	2.19	2.72	1.99	2.63	1.70	2.64
32	1.93	2.64	1.85	2.52	1.75	2.69
33	1.56	2.52	1.66	2.38	1.81	2.74
34	1.28	2.36	1.46	2.25	1.86	2.80
35	1.20	2.21	1.33	2.17	1.90	2.87
36	1.22	2.17	1.31	2.21	1.95	2.95
37	1.38	2.28	1.44	2.40	1.98	3.03
38	1.63	2.74	1.67	2.76	2.02	3.13
39	1.96	3.56	1.96	3.24	2.05	3.24
40	2.22	4.62	2.21	3.75	2.07	3.37
41	2.48	5.61	2.39	4.20	2.09	3.51
42	2.59	6.34	2.45	4.54	2.11	3.67
43	2.59	6.63	2.40	4.69	2.12	3.85
44	2.34	6.40	2.30	4.68	2.13	4.06
45	2.07	5.79	2.20		2.14	4.30
46		5.27	2.17		2.14	4.57
47			2.26		2.15	4.88
48			2.54		2.15	5.24
49			3.05		2.16	5.64
50					2.18	6.11
51					2.20	6.64
52					2.24	7.25
53					2.31	7.95
54					2.40	8.75
55					2.54	9.67
56						10.72
57						11.93

policy years combined. It is generally important to see that the rates for the two sets of years run smoothly and consistently in relation to each other. In order to bring this about, a further adjustment to the rates is often necessary. In table number nineteen, it is seen that Makeham's formula easily gives the most consistent and reasonable results in this respect. It might be well to point out that the rate of disability per thousand in the column headed "Second and Subsequent Years", opposite, say, age 30 at entry, (namely 2.80, by Higham's formula) is that applicable to the second policy year, for policies issued at age 30, or to the third policy year, for policies issued at age 29, and so on. That is, it is an ultimate rate.