WORKMEN'S COMPENSATION D-RATIO REVISIONS

by

Arthur L. Bailey

D-ratios are the ratios of primary losses to total losses where the amount of primary loss corresponding to a specified total loss is as defined in the multi-split experience rating plan. The classification D-ratios are a very important element in the multi-split experience rating process and the experience modification, for a risk having appreciable credibility, is largely dependent upon the D-ratios. Thus it seems necessary that as much care be taken in the revision of classification D-ratios as in the revision of the rates to which such D-ratios are to be applied in the experience rating process. It is the purpose of this paper to present a method for the revision of D-ratios, which utilizes all of the data available and which is designed to produce as accurate D-ratios as possible.

Consideration has recently been given to the use of the primary-excess separation of losses in the revision of rates in lieu of the separation of losses by serious, non-serious and medical. If such a primary-excess separation of losses were to be used in revising rates, the revision of D-ratios might well become an important part of the process of revising rates. For example, if comparatively accurate classification D-ratios could be obtained, the rate revision process might be one whereby the primary pure premiums would be revised on the basis of the available primary loss experience and the total pure premiums would be the revised primary pure premiums divided by the revised D-ratios.

The present paper consists of a description of a process for revising D-ratios together with an appendix describing various functional D-ratios, one of which, based solely on the primary loss experience, is suggested for use in the revision of classification D-ratios.

There are available for the revision of classification D-ratios the following for each classification:

D_u = the underlying D-ratio (namely the D-ratio used prior to the revision being made)

D_f = the functional D-ratio indicated by the average primary loss in the experience to be used for the revision (see Appendix A for a complete description of this functional D-ratio)

D_i = the D-ratio indicated by the experience to be used for the revision (namely the ratio of primary to total losses in such experience)

The first of these presumably summarizes all of the experience available in previous revisions. The second summarizes the indications of the primary losses of the current experience, the primary losses being those which are least subject to chance variation. The third summarizes the indications of all of the current experience including the excess losses which are subject to considerable chance fluctuations.

It is suggested that the revised D-ratios be determined from

$$\text{Revised D-Ratio} = D_r = Z_i D_i + Z_f D_f + (1 - Z_i - Z_f)\, D_u$$

Where Z_i and Z_f are credibilities to be determined with the understanding that $Z_i + Z_f$ is not greater than unity. In other words, the indicated D-ratios will first be given such credibilities as they warrant, the D-ratios indicated by the primary loss portion of the experience will next be given such additional credibilities as they warrant, and the remaining credibilities will be given to the underlying D-ratios.

As the chance variation in D_i will be caused principally by the chance variation in the amount of excess losses, it would seem proper to give Z_i a value equal to the credibility applicable to the amount of excess loss when N losses have occurred. Such a credibility would be of the form:*

$$Z_i = \frac{N}{N + K_e}$$

where N is the total number of losses and K_e is approximated from the excess loss classification experience and data as to the distribution of losses by size of loss.

The chance variation in D_f will be wholly the result of the chance variation in the amount of primary loss and it would seem proper to assign a value to the total of Z_i and Z_f equal to the credibility applicable to the amount of primary loss when N losses have occurred. Such a credibility would be:

$$Z_i + Z_f = \frac{N}{N + K_p}$$

where N is the total number of losses and where K_p is approximated from the primary loss classification experience and data as to the distribution of losses by size of loss.

In actual calculation it will be easier to calculate:

$$Z_u = 1 - Z_i - Z_f = \frac{K_p}{N + K_p}$$

and to obtain Z_f from $Z_f = 1 - Z_i - Z_u$.

On the basis of the Massachusetts D-ratios effective 12/31/45, the Massachusetts classification experience for policy years 1939 to 1943 inclusive, and the distribution of losses by size for policy years 1940 and 1941, the writer approximated the value of K_p as 47 and the value of K_e as 1344.

The entire process of revising the D-ratios using the above procedure is shown in Table I for certain of the Massachusetts classifications while Table II shows for each of the Massachusetts classifications the values of the manual D-ratios effective 12/31/45, the D-ratios indicated by the policy year 1939 to 1943 experience, the functional D-ratios based on the average primary loss in the same experience, and the D-ratios revised under the above procedure.

---

* See "A Generalized Theory of Credibility", P.C.A.S. Vol. XXXII, p. 19.

TABLE I
REVISION OF D-RATIOS
MASSACHUSETTS CLASSIFICATION EXPERIENCE FOR POLICY YEARS 1939 TO 1943 INCLUSIVE

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Number of	Amount of Losses		Average Primary	D-Ratios			Credibilities			
Class	Losses N	Primary	Total	Loss M_p (3)÷(2)	Indicated D_i (3)÷(4)	Functional D_f (Table III)	Under-lying D_u *	Indicated Z_i $\frac{(2)}{(2)+1344}$	Functional Z_f 1.0-(9) -(11)	Underlying Z_u $\frac{47}{(2)+47}$	Revised D-Ratio (6)x(9) +(7)x(10) +(8)x(11)
2070	1,430	$293,402	$370,976	$205	.79	.77	.73	.52	.45	.03	.78
2089	790	104,129	115,738	132	.90	.86	.87	.37	.57	.06	.88
2105	15	5,338	5,496	356	.97	.59	.83	.01	.23	.76	.78
2110	24	5,352	8,639	223	.62	.75	.74	.02	.32	.66	.74
2121	486	110,658	153,217	228	.72	.74	.78	.27	.64	.09	.74
2291	544	120,170	140,093	221	.86	.75	.72	.29	.63	.08	.78
2586	312	59,557	64,518	191	.92	.79	.83	.19	.68	.13	.82
2660	3,564	785,965	963,609	221	.82	.75	.79	.73	.26	.01	.80
2686	23	4,115	4,115	179	1.00	.80	.80	.02	.31	.67	.80
3515	476	116,801	150,388	245	.78	.72	.65	.26	.65	.09	.73
3516	15	3,352	3,353	223	1.00	.75	.75	.01	.23	.76	.75
3559	6	3,039	3,182	507	.96	.40	.79	-	.11	.89	.75
3632	4,751	1,204,871	1,544,496	254	.78	.71	.79	.78	.21	.01	.77
4362	10	4,781	8,589	478	.56	.44	.67	.01	.17	.82	.63
5348	54	15,707	32,689	291	.48	.67	.70	.04	.49	.47	.68
5403	1,225	361,153	573,703	295	.63	.66	.62	.48	.48	.04	.64
5551	445	117,864	189,887	265	.62	.70	.50	.25	.65	.10	.66
8018	1,303	285,081	372,186	219	.77	.75	.80	.49	.48	.03	.76
8044	458	98,789	170,080	216	.58	.76	.75	.25	.66	.09	.71
9052	1,376	302,279	434,364	220	.70	.75	.61	.51	.46	.03	.72
9079	4,062	738,709	904,509	182	.82	.80	.78	.75	.24	.01	.81
9170	35	12,039	22,956	344	.52	.60	.73	.03	.40	.57	.67

*From Experience Rating Plan - D-Ratios effective December 31, 1945

TABLE II

COMPARISON OF D-RATIOS FOR MASSACHUSETTS CLASSIFICATIONS

D_u = Underlying D-Ratio from the Experience Rating Plan effective 12/31/45

D_i = Indicated D-Ratio from Massachusetts experience for Policy Years 1939-1943 inclusive

D_f = Functional D-Ratio corresponding to Average Primary Loss for Policy Years 1939-1943 inclusive

D_r = Revised D-Ratio calculated as shown in Table I

Class	D_u	D_i	D_f	D_r	Class	D_u	D_i	D_f	D_r	Class	D_u	D_i	D_f	D_r
0006	.73	.67	.75	.72	2173	.78	1.00	.88	.78	2503	.77	.91	.67	.75
0042	.75	.88	.79	.79	2174	.90	.98	.84	.88	2531	.84	.96	.75	.81
0912	.65	.69	.69	.69	2177	.84	.94	.77	.82	2532	.81	.89	.78	.80
1748	.65	.68	.79	.75	2211	.61	.63	.71	.66	2534	.74	1.00	.80	.74
2001	.67	.77	.75	.74	2216	.61	.61	.70	.67	2537	.75	.99	.79	.77
2002	.81	.75	.67	.75	2220	.79	.79	.77	.78	2538	.81	.69	.77	.77
2003	.79	.76	.76	.76	2222	.79	.78	.78	.78	2560	.84	.83	.46	.82
2014	.79	.82	.73	.76	2260	.71	.70	.72	.71	2570	.75	.79	.77	.77
2016	.71	.97	.68	.71	2280	.90	1.00	.87	.90	2571	.78	.86	.73	.77
2021	.79	.75	.79	.78	2286	.72	.72	.75	.73	2575	.78	.60	.72	.75
2039	.87	.73	.73	.76	2288	.69	.74	.75	.74	2576	.74	.74	.69	.71
2041	.82	.78	.77	.78	2291	.72	.86	.75	.78	2578	.77	.95	.74	.77
2042	.81	.79	.65	.70	2300	.73	.74	.71	.72	2585	.72	.70	.76	.73
2045	.82	.39	.73	.78	2302	.83	.67	.75	.77	2586	.83	.92	.79	.82
2070	.73	.79	.77	.78	2303	.83	.81	.75	.77	2587	.72	.97	.80	.75
2081	.83	.67	.76	.78	2348	.72	.66	.74	.73	2600	.85	.97	.46	.84
2089	.87	.90	.86	.88	2351	.79	.85	.84	.82	2610	.77	1.00	.85	.78
2095	.79	.72	.77	.76	2352	.76	.69	.68	.70	2623	.72	.73	.75	.74
2101	.90	.72	.82	.82	2361	.78	.93	.80	.81	2640	.73	.93	.76	.75
2105	.83	.97	.59	.78	2362	.82	.81	.75	.77	2651	.75	.77	.73	.75
2110	.74	.62	.75	.74	2380	.74	.64	.79	.76	2654	.75	.61	.71	.73
2111	.81	.81	.80	.80	2384	.74	.91	.79	.75	2660	.79	.82	.75	.80
2112	.80	.82	.74	.77	2386	.85	1.00	.91	.85	2670	.79	1.00	.82	.79
2114	.84	1.00	.93	.85	2387	.74	.64	.80	.77	2681	.75	1.00	.86	.75
2121	.78	.72	.74	.74	2388	.78	1.00	.80	.78	2683	.82	.85	.82	.82
2131	.80	.95	.79	.80	2402	.81	.75	.74	.75	2686	.80	1.00	.80	.80
2143	.76	.52	.69	.73	2413	.82	.79	.74	.76	2688	.85	.85	.75	.78
2156	.80	.98	.77	.80	2416	.78	.81	.79	.79	2795	.83	.98	.75	.80
2157	.78	.78	.78	.78	2417	.80	.76	.74	.75	2812	.73	.71	.65	.68
2163	.78	.90	.77	.78	2501	.78	.82	.79	.81	3042	.81	.81	.71	.76

TABLE II (cont'd)

Class	D_u	D_i	D_f	D_r	Class	D_u	D_i	D_f	D_r	Class	D_u	D_i	D_f	D_r
3060	.85	.76	.55	.76	5146	.63	.96	.80	.70	7500	.72	.69	.68	.69
3066	.86	.79	.73	.76	5160	.64	.56	.63	.63	7502	.68	1.00	.89	.68
6076	.81	.75	.71	.73	5183	.72	.72	.72	.72	7539	.56	.53	.63	.59
3152	.79	.77	.70	.73	5184	.72	.80	.67	.70	7570	.69	1.00	.81	.69
3188	.84	.58	.73	.77	5188	.68	.60	.59	.64	7600	.56	1.00	.89	.60
3200	.84	.84	.71	.76	5190	.57	.55	.65	.62	7609	.57	.96	.56	.57
3315	.77	.78	.73	.74	5200	.82	.67	.72	.73	8006	.87	.80	.76	.78
3316	.77	.68	.69	.72	5203	.60	.62	.69	.63	8007	.76	.69	.81	.79
3400	.73	.77	.69	.72	5213	.62	.69	.62	.65	8017	.80	.79	.76	.78
3507	.80	.73	.77	.78	5215	.62	.55	.66	.63	8018	.80	.77	.75	.76
3515	.65	.78	.72	.73	5348	.70	.48	.67	.68	8032	.78	.78	.76	.77
3516	.75	1.00	.75	.75	5403	.62	.63	.66	.64	8044	.75	.58	.76	.71
3527	.80	.71	.73	.75	5437	.63	.62	.67	.66	8049	.80	.93	.82	.81
3548	.79	.96	.73	.77	5443	.69	.88	.77	.71	8090	.80	.99	.73	.79
3558	.85	.80	.60	.67	5461	.63	.50	.56	.56	8103	.77	.83	.73	.75
3559	.79	.96	.40	.75	5462	.65	.73	.71	.69	8105	.75	.73	.75	.75
3561	.74	.96	.65	.73	5480	.69	.66	.67	.67	8264	.84	.75	.80	.80
3574	.84	.95	.69	.78	5490	.53	.62	.66	.64	8285	.83	1.00	.70	.82
3612	.78	.77	.72	.74	5491	.53	1.00	.85	.54	8295	.77	.87	.76	.77
3629	.84	.78	.71	.77	5538	.62	.72	.71	.70	8731	.71	1.00	.42	.70
3632	.79	.78	.71	.77	5551	.50	.62	.70	.66	8745	.84	.97	.76	.80
3634	.71	.75	.70	.72	5645	.63	.61	.69	.65	8833	.75	.63	.71	.70
3635	.90	.89	.67	.74	5651	.63	.78	.71	.70	9015	.71	.64	.70	.67
8639	.76	.75	.70	.72	6204	.69	.69	.69	.69	9052	.61	.70	.75	.72
3724	.65	.63	.65	.65	6504	.84	.78	.73	.76	9053	.84	.64	.68	.82
3726	.71	.68	.58	.68	6824	.69	.81	.67	.70	9060	.83	.89	.78	.80
3305	.78	1.00	.76	.78	6872	.61	.65	.67	.65	9061	.85	.79	.77	.79
4000	.64	.61	.65	.64	7201	.71	1.00	.77	.73	9063	.79	.76	.75	.76
4273	.86	.98	.80	.83	7205	.68	.83	.75	.71	9079	.78	.82	.80	.81
4279	.69	.67	.74	.71	7207	.73	.97	.81	.75	9090	.75	.73	.84	.81
4562	.67	.56	.44	.63	7219	.74	.73	.75	.74	9170	.73	.52	.60	.67
4501	.84	.69	.54	.82	7309	.68	.71	.76	.74	9800	.63	.66	.56	.59
5000	.62	.76	.63	.62	7380	.71	.67	.75	.71					
5022	.62	.57	.65	.62	7382	.69	.66	.73	.71					
5059	.48	.65	.57	.50	7392	.79	.78	.77	.78					

APPENDIX A

FUNCTIONAL D-RATIO

For each classification there exists an expected or true value of the ratio of primary losses to total losses which we shall indicate by the script $\mathcal{D}$. From classification experience, there will be available an "indicated" ratio of primary losses to total losses which we shall designate as D. In addition to the indicated D-ratios there are available what we shall call functional D-ratios which are the estimates of $\mathcal{D}$ made from observed value of other statistics such as:

D_1 determined from $\frac{1}{M_p}$ where M_p is the average amount of primary loss per loss.

D_2 " " R which is the ratio of the number of losses which involve an excess loss to the total number of losses.

D_3 " " $\frac{R}{M_p}$ which is the reciprocal of the average amount of primary loss per loss which involves an excess loss.

D_4 " " M_e which is the average amount of excess loss per loss which involves an excess loss.

D_5 " " M'_e which is the average amount of excess loss per loss.

D_6 " " M_e/M_p = the ratio of M_e to M_p.

We need to determine whether any of these functional D-Ratios are more efficient estimates of the true $\mathcal{D}$ than is the indicated D, and if so, which one is the most efficient.

In order to obtain results which could be directly assimilated without undue use of mathematical formulae, it seemed best to make an empirical test. This consisted of preparing 1,000 punch cards each representing a single loss such that the 1,000 losses have a distribution by size of loss approximately the same as that for all losses in Massachusetts in 1940-41. Random numbers were assigned to the 1,000 punch cards and 100 random samples of 50 losses each were thus obtained. The indicated values of each of the statistics mentioned above were calculated for each of the 100 observations and the squares of the coefficients of variation of these 100 observations were calculated for each statistic.

It will be realized that in this way we have made 100 empirical observations of experience for a single classification for which the true values are known and are thus in a position to establish which of the statistics are stable and which are subject to wide chance variation. The squares of the coefficient of variation were as follows:

Statistic	Square of Coefficient of Variation
$\frac{1}{D} - 1$	.6793
$\frac{1}{M_p}$	.0466
R	.1443
$\frac{R}{M_p}$	.0681
M_e	1.1021
M'_e	.8234
$\frac{M_e}{M_p}$	1.3927

The statistics were handled in the above forms in order that they might be on as nearly a uniform basis as possible. The following functional relationships between these statistics will indicate why the above forms were chosen:

$$\left(\frac{1}{D} - 1\right) = \left(\frac{1}{M_p}\right)(M'_e) = (R)\left(\frac{M_e}{M_p}\right) = \left(\frac{R}{M_p}\right)(M_e)$$

The comparatively small sampling variation in $\frac{1}{M_p}$, R and $\frac{R}{M_p}$ would indicate the possibility that these statistics might produce more efficient estimates of $\mathcal{D}$ than the indicated values of D.

The experience for all of the Massachusetts classifications were cut on punch cards together with the values of these three statistics calculated from the classification experience. These classification cards were then sorted successively by each of these statistics and tabulated to obtain approximately the relationships between $\left(\frac{1}{\mathcal{D}} - 1\right)$ and $\frac{1}{M_p}$, $\mathcal{R}$ and $\frac{\mathcal{R}}{M_p}$ (where the script letters indicate the true values as contrasted to sample values). Charts were then drawn for each of the three statistics similar to that shown in Figure A for $\left(\frac{1}{\mathcal{D}} - 1\right)$ and $\frac{1}{M_p}$.

By entering these charts with the observed values of $\frac{1}{M_p}$, R and $\frac{R}{M_p}$ obtained from the 100 random samples, the functional estimates of $(\frac{1}{\mathcal{D}} - 1)$ were obtained for comparison as to efficiency with the observed values of $(\frac{1}{D} - 1)$.

The following squares of coefficients of variation were obtained:

Statistic	Squares of Coefficients of Variation
$(\frac{1}{\mathcal{D}} - 1)$ estimated from $(\frac{1}{D} - 1)$	.6793
$(\frac{1}{\mathcal{D}} - 1)$ estimated from $\frac{1}{M_p}$	.1007
$(\frac{1}{\mathcal{D}} - 1)$ estimated from R	.1443
$(\frac{1}{\mathcal{D}} - 1)$ estimated from $\frac{R}{M_p}$	.1824

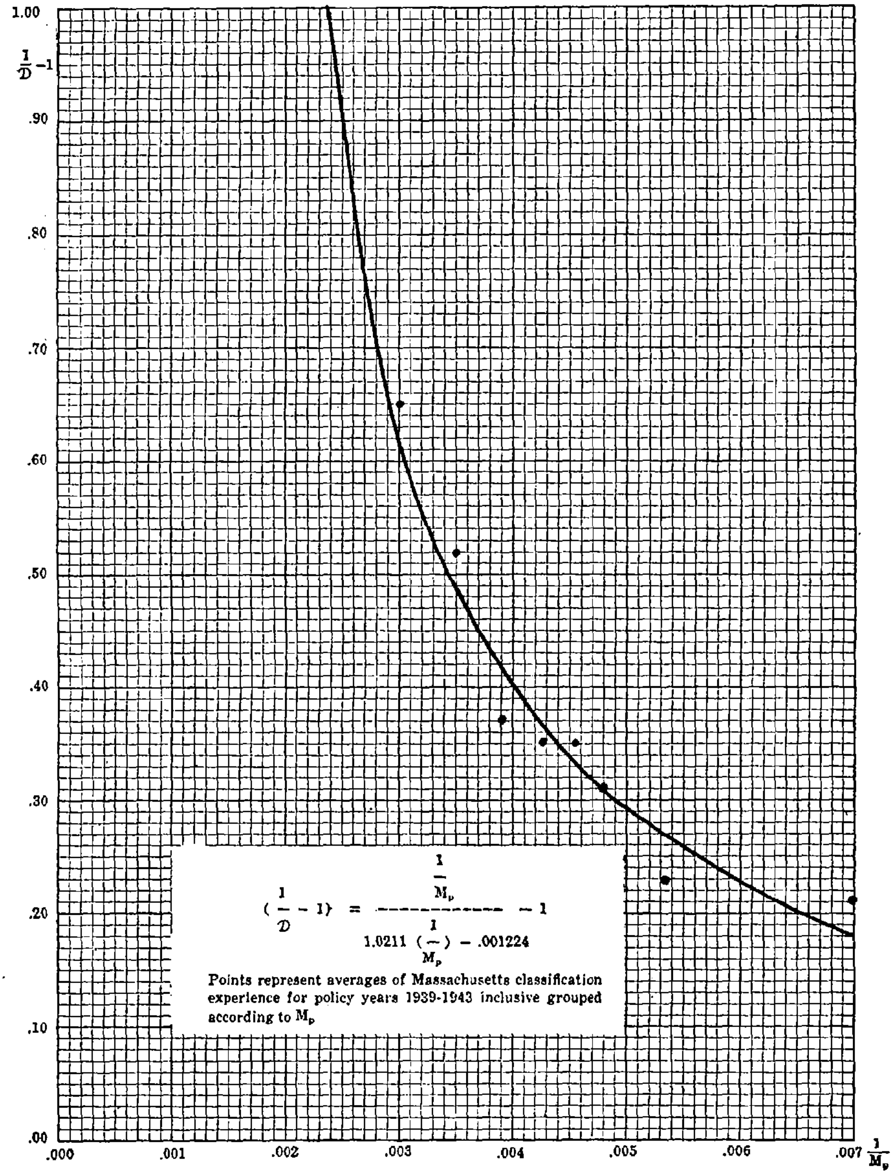
Thus we find that each of these three statistics is more efficient than $(\frac{1}{D} - 1)$ itself and that $\frac{1}{M_p}$ appears to be the most efficient. It seemed worth while, however, to substantiate this with samples from distributions of losses having rather different values of $\mathcal{D}$. To do this the losses on the 1,000 punch cards representing the average Massachusetts distribution of losses were first reduced 50% and then increased 50% and the new values of primary and excess losses determined. Twenty samples of 50 losses each were then taken from each of these new empirical classifications and the estimated values of $(\frac{1}{\mathcal{D}} - 1)$ determined from the charts as before. The squares of the coefficient of variation were as follows:

$(\frac{1}{\mathcal{D}} - 1)$ Estimated From	Squares of Coefficient of Variation: Losses Increased 50% (1)	Average Loss Distribution (2)	Losses Reduced 50% (3)	Weighted Averaged of (1), (2) and (3)
$(\frac{1}{\mathcal{D}} - 1)$	.4431	.6793	1.2800	.7314
$\frac{1}{M_p}$	.0738	.1007	.1004	.0968
R	.1434	.1443	.1880	.1504
$\frac{R}{M_p}$	.1169	.1824	.2091	.1769

It is seen that the estimates made on the basis of the value of M_p are the most efficient in each case.

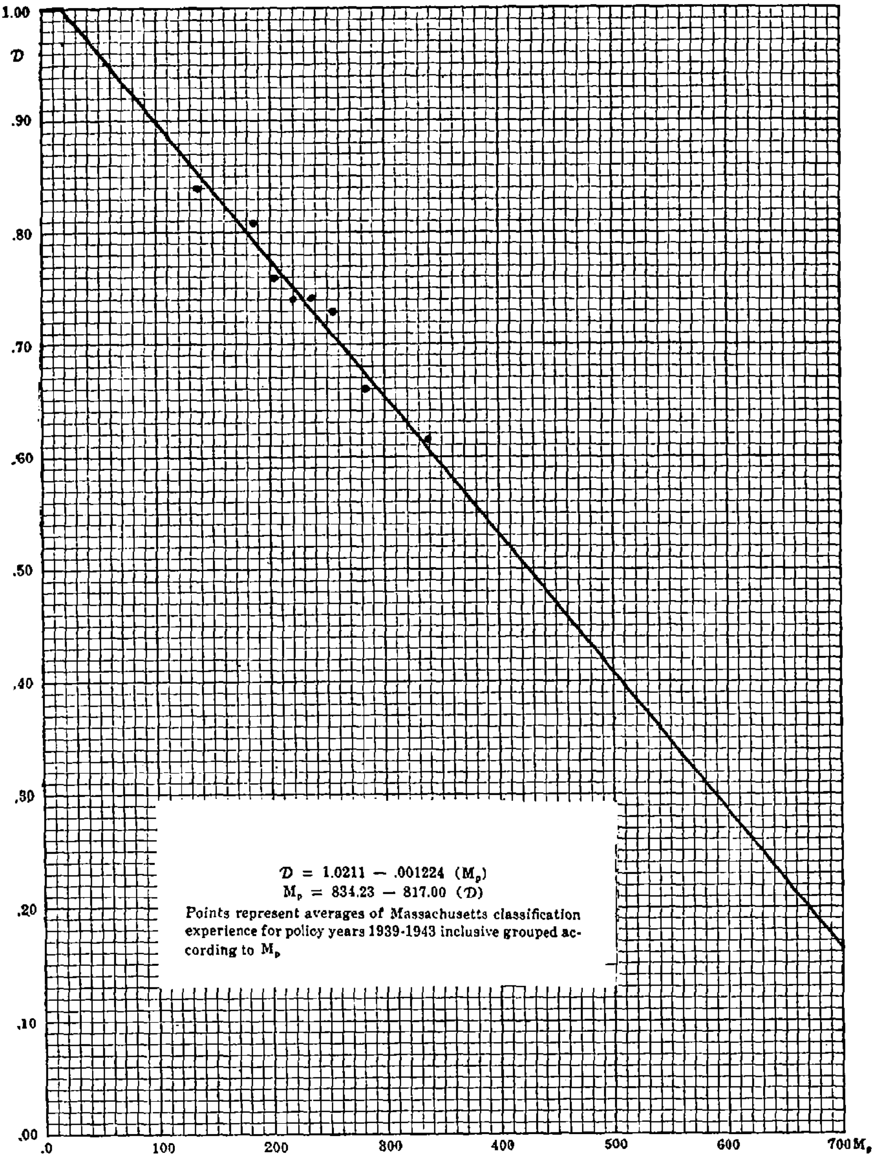
Figure B and Table III, expressing $\mathcal{D}$ in terms of M_p were obtained from Figure A and provide direct means for estimating $\mathcal{D}$ from M_p. Figure C shows a comparison of the true $\mathcal{D}$-Ratios with the indicated values of D for each of the three empirical classifications and Figure D shows the same comparison of the true $\mathcal{D}$-Ratios with the functional D-Ratios estimated from the values of M_p.

FIGURE A



1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
.00
$\frac{1}{D}-1$
.000
.001
.002
.003
.004
.005
.006
.007
$\frac{1}{M_p}$
$$\left(\frac{1}{D}-1\right) = \frac{\frac{1}{M_p}}{1.0211\left(\frac{1}{M_p}\right) - .001224} - 1$$
Points represent averages of Massachusetts classification experience for policy years 1939-1943 inclusive grouped according to M_p

FIGURE B



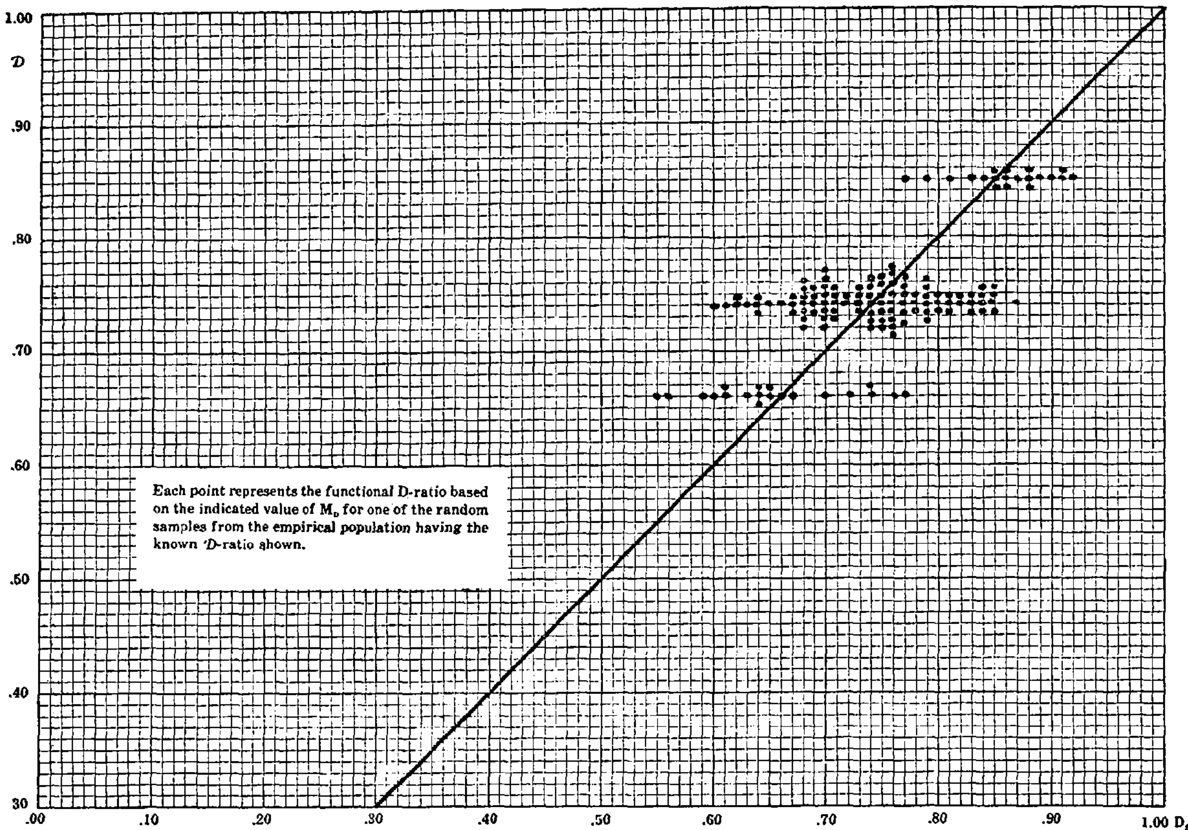
1.00
𝒟
.90
.80
.70
.60
.50
.40
.30
.20
.10
.00
.0
100
200
300
400
500
600
700 M_p
$\mathcal{D} = 1.0211 - .001224\ (M_p)$
$M_p = 834.23 - 817.00\ (\mathcal{D})$
Points represent averages of Massachusetts classification experience for policy years 1939-1943 inclusive grouped according to M_p

TABLE III

FUNCTIONAL D-RATIOS, D_f, FOR VARIOUS AVERAGE PRIMARY LOSSES

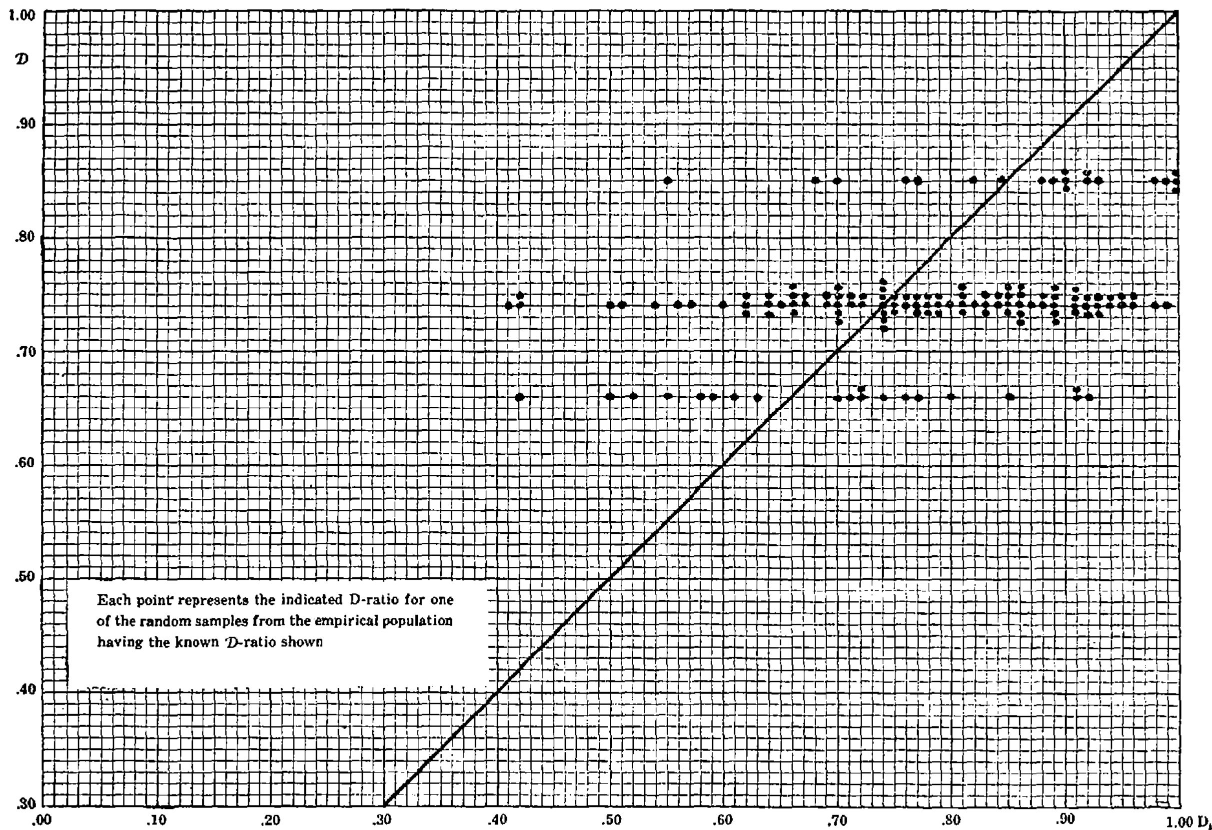
M_p	D_f	M_p	D_f
0 - 21	1.00	341 - 348	.60
22 - 29	.99	349 - 356	.59
30 - 38	.98	357 - 364	.58
39 - 46	.97	365 - 373	.57
47 - 54	.96	374 - 381	.56
55 - 62	.95	382 - 389	.55
63 - 70	.94	390 - 397	.54
71 - 79	.93	398 - 405	.53
80 - 87	.92	406 - 413	.52
88 - 95	.91	414 - 422	.51
96 - 103	.90	423 - 430	.50
104 - 111	.89	431 - 438	.49
112 - 119	.88	439 - 446	.48
120 - 128	.87	447 - 454	.47
129 - 136	.86	455 - 462	.46
137 - 144	.85	463 - 471	.45
145 - 152	.84	472 - 479	.44
153 - 160	.83	480 - 487	.43
161 - 168	.82	488 - 495	.42
169 - 177	.81	496 - 503	.41
178 - 185	.80	504 - 512	.40
186 - 193	.79	513 - 520	.39
194 - 201	.78	521 - 528	.38
202 - 209	.77	529 - 536	.37
210 - 217	.76	537 - 544	.36
218 - 226	.75	545 - 552	.35
227 - 234	.74	553 - 561	.34
235 - 242	.73	562 - 569	.33
243 - 250	.72	570 - 577	.32
251 - 258	.71	578 - 585	.31
259 - 266	.70	586 - 593	.30
267 - 275	.69	594 - 601	.29
276 - 283	.68	602 - 610	.28
284 - 291	.67	611 - 618	.27
292 - 299	.66	619 - 626	.26
300 - 307	.65	627 - 634	.25
308 - 315	.64	635 - 642	.24
316 - 324	.63	643 - 650	.23
325 - 332	.62	651 - 659	.22
333 - 340	.61	660 - 667	.21

FIGURE D



1.00
𝒟
.90
.80
.70
.60
.50
.40
30
.00
.10
.20
.30
.40
.50
.60
.70
.80
.90
1.00 D_t
Each point represents the functional D-ratio based on the indicated value of M_D for one of the random samples from the empirical population having the known 𝒟-ratio shown.

FIGURE C


1.00
𝒟
.90
.80
.70
.60
.50
.40
.30
.00
.10
.20
.30
.40
.50
.60
.70
.80
.90
1.00 D_i
Each point represents the indicated D-ratio for one of the random samples from the empirical population having the known 𝒟-ratio shown