Biometrika (1979), **66**, 1, *pp.* 191–3
Printed in Great Britain

A note on the intervals between coal-mining disasters

BY R. G. JARRETT
C.S.I.R.O. Division of Mathematics & Statistics, Melbourne

SUMMARY

The data given by Maguire, Pearson & Wynn (1952) on the time intervals between successive coal-mining disasters involving 10 or more men killed have been found to contain a number of errors. The original numbers have been quoted and used by a number of authors. The present note corrects the errors in the original paper and extends the data set to cover 191 disasters between 1851 and 1962.

Some key words: Interval; Point process; Poisson process; Series of events; Trend.

1. INTRODUCTION

During practical sessions for an M.Sc. course in Statistics at Imperial College, the data of Maguire, Pearson & Wynn (1952) on the intervals between coal-mining disasters were used as an example. It was noticed that the time interval between 6 December 1875 and 29 May 1951, the starting and finishing dates, did not tally with the sum of the 109 intervals, namely 26,263 days. Since these data have been extensively used in the literature, it was thought worthwhile to place the correct values on record.

Table 1. *Time intervals in days between explosions in mines, from* 15 *March* 1851 *to* 22 *March* 1962 (*to be read down columns*)

157	65	53	93	127	176	22	1205	1643	312
123	186	17	24	218	55	61	644	54	536
2	23	538	91	2	93	78	467	326	145
124	92	187	143	0	59	99	871	1312	75
12	197	34	16	378	315	326	48	348	364
4	431	101	27	36	59	275	123	745	37
10	16	41	144	15	61	54	456	217	19
216	154	139	45	31	1	217	498	120	156
80	95	42	6	215	13	113	49	275	47
12	25	1	208	11	189	32	131	20	129
33	19	250	29	137	345	388	182	66	1630
66	78	80	112	4	20	151	255	292	29
232	202	3	43	15	81	361	194	4	217
826	36	324	193	72	286	312	224	368	7
40	110	56	134	96	114	354	566	307	18
12	276	31	420	124	108	307	462	336	1358
29	16	96	95	50	188	275	228	19	2366
190	88	70	125	120	233	78	806	329	952
97	225	41	34	203	28	17	517	330	632

The original data were traced to the *Colliery Year Book and Coal Trades Directory*, available from the National Coal Board in London. Publication ceased in 1962, but issues before then give data on explosions in coal mines back to 15 March 1851. Accordingly, the data have been corrected and extended to cover the period 15 March 1851 to 22 March 1962 inclusive, a total of 40,550 days. There were 191 explosions involving 10 or more men killed, including

ones on each of the above dates. The 190 intervals between explosions shown in Table 1 therefore have a sum of 40,549 days. The zero in Table 1 occurs because there were two accidents on the same day on 6 December 1875, the date used by Maguire, Pearson & Wynn as their starting point. Prior to 1900, the *Colliery Year Book* lists 'disasters caused by explosions of fire-damp or coaldust', whereas subsequent data gives 'principal disasters from all causes', explosions being one of several subheadings. The data tabulated here include only those accidents since 1900 specifically labelled as explosions.

2. Some calculations

The data given in the original paper have been used extensively in the literature to illustrate various methods of handling point processes; see, for example, Barnard (1953), Cox & Lewis (1966), Boneva, Kendall & Stefanov (1971). We give here just a few simple calculations to show that the main conclusions about the data are unchanged.

Table 2 shows the distribution of disasters over the days of the week and the months of the year. Noticeably fewer disasters occurred on Sundays and Mondays. If the disasters were randomly distributed over the days of the year, the number of days having zero, one, two, ... disasters would be Poisson with mean 0·523. Table 3 shows good agreement between observed and expected values. The assumption of exponentially distributed intervals would imply a Poisson process with rate 190/40,549 disasters per day, whence an expected number of 0·444 days would have two or more disasters. The occurrence of two disasters on one day, as was observed, is therefore reasonable.

Table 2. *Distribution of disasters over days of the week and months of the year*

Day of week	Sun.	Mon.	Tues.	Wed.	Thurs.	Fri.	Sat.
Number	5	19	34	33	36	35	29

Month	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Number	14	20	20	13	14	10	18	15	11	16	16	24

Table 3. *Poisson distribution fitted to number of disasters on different days of the year, excluding February* 29

No. of disasters	0	1	2	3	>3
No. of days with this number of disasters	214	117	28	6	0
Fitted number of days	216·3	113·2	29·6	5·2	0·7

Table 4. *Analysis of variance of log interval between successive sets of* 5 *accidents on time from origin*

	Degrees of freedom	Mean square
Linear, ignoring quadratic	1	9·543
Quadratic, adjusting for linear	1	0·002
Residual	35	0·309
Theoretical residual	—	0·222

Figure 1 shows a plot of the cumulative number of events against the number of days after 15 March 1851. This may be compared with Fig. 1.1 of Cox & Lewis (1966, p. 3), their zero corresponding to event number 81 and a time of 9032 days in the present graph. As Cox & Lewis (1966, p. 42) showed, the gradual decrease in the accident rate may be illustrated by a quadratic regression of $\log(y/100)$, where y is the interval between each successive set of 5 events, on the time z from the origin to the midpoint of the corresponding interval. Table 4

shows the analysis of variance table together with the theoretical residual. The quadratic term is not significant and the fitted equation is

$$\log(y/100) = 2{\cdot}07 + 0{\cdot}476(z - 13{,}763) \times 10^{-4}.$$

The plot in Fig. 1 appears to be quite linear up to about 125 events. The above test applied to the first 125 intervals gives a mean rate of one event every 106 days with insignificant linear and quadratic terms, while the rate for the next 65 intervals is about one every 338 days. An analysis involving man-hours worked or tons of coal produced may reveal more about the reasons for such changes in the behaviour of the process.

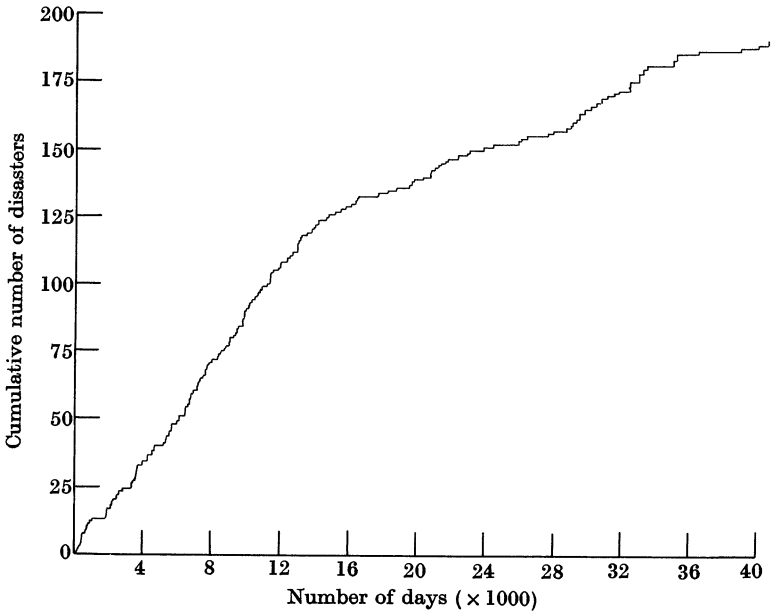

Fig. 1. Cumulative number of disasters against time.

References

Barnard, G. A. (1953). Time intervals between accidents—a note on Maguire, Pearson and Wynn's paper. *Biometrika* **40**, 212–3.

Boneva, L. I., Kendall, D. G. & Stefanov, I. (1971). Spline transformations: three new diagnostic aids for the statistical data-analyst (with discussion). *J. R. Statist. Soc.* B **33**, 1–71.

Cox, D. R. & Lewis, P. A. W. (1966). *The Statistical Analysis of Series of Events*. London: Methuen.

Maguire, B. A., Pearson, E. S. & Wynn, A. H. A. (1952). The time intervals between industrial accidents. *Biometrika* **39**, 168–80.

[*Received October* 1978]