

Shorter contribution

The effect of rain on attendance at Sydney's Easter Show: an example of the double normalisation of data

Edward Linacre

School of Earth Sciences, Macquarie University, Australia

(Manuscript received July 1989; revised August 1989)

The following is an example of double normalisation, a simple procedure which does not appear to have been described before. It allows useful deductions from data that at first sight appear hopelessly confused by variations of two distinct sorts, together. The procedure could have wide application. In the present case, the variations are due to day-to-day fluctuations and, separately, to trends over several years.

The example concerns a major annual event in Sydney, the agricultural show, which has been held around Eastertime for many years. It starts on the Friday before Good Friday, and finishes on the Tuesday after Easter Sunday. The show was not open on Easter Sunday itself, prior to 1983.

The present exercise is aimed at examining the connection between daily attendance and the occurrence of rain during the day (i.e. measured in the morning of the following day). The raw data are shown in Table 1. The bottom row shows a rise to maximum attendances on Good Friday and the day after, and the last column shows some trend

towards rising attendances until 1981, followed by a decline.

Before it is possible to examine the effects of rainfall on attendance, the influence of the two variations (i.e. within each show period, and between the periods) has to be removed. This is done in the next tables. Table 2 illustrates normalisation according to the day of the week, by dividing each figure in Table 1 by the mean for that day of the week. Table 3 shows the normalisation of the year-to-year trend, by dividing each figure in Table 2 by the three-year running-mean of the ratio of the annual attendance to the overall mean annual attendance. This device was adopted to remove the effect of any general trend without removing the effect of any particularly wet year. Choosing three years as the period of averaging removes any biennial oscillation, but avoids reducing the number of rows of data unduly.

The last column in Table 3 shows that factors other than year-to-year and day-to-day trends lead to a scatter of the 12-day attendance, equivalent to

Table 1. Raw data on the attendance at the Easter Show, in thousands each day.

Year	Day in each year							Fri	Sat	ES*	Mon	Tue	Total
	Fri	Sat	Sun	Mon	Tue	Wed	Thu						
1975	33.9	111	52.2	27.5	72.4	51.6	69.4	145	159		131	50.9	903
1976	33.8	95.3	53.0	25.2	50.3	51.4	64.6	139	169		106	55.5	843
1977	33.0	101	62.2	30.5	56.1	53.9	68.3	164	171		139	69.2	949
1978	28.5	52.4	23.8	15.5	36.3	56.3	74.9	166	175		165	74.5	868
1979	33.7	107	78.5	33.8	57.4	62.0	81.9	193	194		141	71.8	1054
1980	33.2	113	88.5	40.5	59.6	72.4	91.1	184	188		145	62.8	1079
1981	34.1	117	93.4	42.6	65.2	72.6	97.9	206	206		156	72.4	1163
1982	38.8	116	90.2	46.7	59.5	72.3	100	188	199		154	62.5	1127
1983	34.5	103	74.1	40.6	65.2	69.3	95.5	167	157	53.7	109	63.9	1033
1984	35.3	101	82.2	41.4	56.6	65.1	66.9	97	185	83.5	151	73.1	1038
1985	38.3	106	81.5	40.7	57.9	57.4	84.4	175	159	80.5	117	57.7	1055
1986	31.3	87.1	65.1	37.7	57.9	61.5	78.1	156	127	76.3	95	58.0	930
1987	32.6	77.1	67.2	42.5	54.5	58.2	58.8	135	89	81.8	106	51.0	854
1988	29.9	73.4	66.2	36.9	47.9	66.0	58.8	117	111	64.4	71	61.2	803
Mean	33.6	97.2	69.9	35.9	56.9	62.1	77.9	159	163	73.4	128	63.2	979

*i.e. Easter Sunday

Table 2. Part of Table 1, but normalised for the day of the week, i.e. each number in Table 1 has been divided by the mean in the bottom row. All values have been multiplied by 1000. The last column shows the three-year running means of the penultimate column, e.g. 917 is the average of 922, 861 and 969.

Year	Day in each year												Total	
	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	ES	Mon	Tue		
1975	1009	1142	747	766	1272	831	891	912	975	1038	805	922	917	
1976	1006	980	758	702	884	828	829	874	1037	828	878	861		
1977	982	1039	890	850	986	868	877	1031	1049*	1086	1095	969		906
1978	848	539	340	432	638	907	961	1044	1074	1289	1179	887		978

etc.

*For example, this number was calculated from figures in the same column in Table 1, i.e. $1000 \times (171/163)$

Table 3. The 'relative attendance' at the Easter Show, i.e. Table 2 in full, but normalised for year-by-year trends by dividing all numbers by the values in the last column of that table. Again, all numbers have been multiplied by 1000. Italic numbers indicate less than 1 mm of rain, underlining more than 1 mm but less than 5 mm; double underlining means more than that but less than 10 mm, and double underlining and italic shows higher rainfalls.

Year	Day in each year												Total
	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	ES	Mon	Tue	
1976	1097	1069	<u>827</u>	<u>766</u>	964	903	904	953	1131*		<u>903</u>	957	939
1977	1084	1147	<u>982</u>	<u>938</u>	1088	<u>958</u>	<u>968</u>	1138	1158		1199	1209	1070
1978	<i>867</i>	<i>551</i>	<u>348</u>	<u>441</u>	<u>652</u>	<u>927</u>	<u>983</u>	<u>1067</u>	<u>1098</u>		1318	1206	907
1979	981	<u>1077</u>	<u>1099</u>	<u>921</u>	987	977	<u>1028</u>	<u>1188</u>	<u>1164</u>		<u>1078</u>	<u>1112</u>	1054
1980	881	1037	1128	1005	933	1039	1042	1031	1028		1010	886	982
1981	885	1050	1165	1035	999	1019	1096	1130	1102		1063	999	1036
1982	1021	1055	<u>1141</u>	1150	925	1029	<u>1135</u>	<u>1045</u>	<u>1080</u>		<u>1064</u>	874	1018
1983	944	974	<u>974</u>	1040	1053	1026	<u>1127</u>	<u>965</u>	<u>885</u>	<u>673</u>	<u>783</u>	929	970
1984	<u>988</u>	977	<u>1105</u>	<u>1084</u>	935	985	<u>807</u>	<u>573</u>	1067	<u>1070</u>	<u>1109</u>	1087	996
1985	<i>1108</i>	1060	1133	1102	989	<u>898</u>	<u>1052</u>	1070	948	1066	<u>888</u>	<u>887</u>	1048
1986	964	<u>927</u>	963	1086	1053	<u>1023</u>	1036	<u>1014</u>	<u>806</u>	<u>1075</u>	<u>767</u>	949	982
1987	<u>1101</u>	900	1091	1344	1087	<u>1064</u>	857	964	<u>620</u>	1264	940	916	990

*For example, this figure was calculated from numbers in the second row of Table 2, by dividing 1037 by 917, and then multiplying by 1000.

a standard deviation of about 5 per cent of the average. The weather is presumably the reason for the scatter, in view of the effect of rainfall on attendance, seen by the following examination of the Table. The numbers in Table 3 are coded according to the occurrence of rain at Observatory Hill on that day (i.e. collected at 9 am on the following day).

The relative attendance over the 84 days without rain averages 1.036, with a standard deviation of 0.099. The relative attendance is effectively the same on the 19 days when rainfalls were less than 1 mm, i.e. about 1.018 (standard deviation of 0.087). Likewise on the 15 days when the rainfall was between 1-5 mm, the figure is 1.002 (0.152). On the 10 days with 5-10 mm the relative attendance falls to 0.853 (0.172), and on 9 days with more than 10 mm to 0.748 (0.266). These figures imply around 2.5%/mm as the reduction of

attendance for each millimetre of precipitation, at least for rainfalls up to 10 mm or so.

From the point of view of taking out insurance against the reduced attendance due to rain, the important figure is the expected decrease. This can be derived from the figure of 1.036 for the attendance on dry days relative to the average. The inverse, i.e. 0.965, shows that rain reduces attendance by 3.5% on the whole.

However, the sequences in Table 3 reveal that there is some tendency for rain to postpone attendance, rather than reduce it altogether. Thus, the run of bad weather at the start of the show in 1978 led to record relative attendances on the closing two days. Such information is much easier to extract from the normalised values in Table 3 than from the raw data in Table 1.

It is interesting to speculate on what influences attendance. How does the fact that rain is going to

fall later during the day affect the decision about setting out for the show? Perhaps it is the cloudiness of the sky at 9 am, or else some personal estimation of the weather ahead based on the rainfall during the previous day? Simple analysis of Table 3 rules out that last possibility. As regards cloud, the average normalised attendance was 1.058 when there was none, with a standard deviation of 13 values equal to 0.074; when 3 oktas, 0.986 with 0.066 for 9 values; when 5 oktas, 1.000 with 0.074 for 14; when 8 oktas, 0.803 with 0.246 for 15. It seems as though any blue in the sky at 9

am is enough to encourage attendance. But the large standard deviation for an overcast sky suggests some other factor was then operative, in addition. Presumably the occurrence or absence of rain.

Acknowledgments

I am obliged to the Royal Agricultural Society of New South Wales for the attendance figures, and to a referee for helpful comments.

