Click on the **Define event** button, enter 1, and click on **Continue**, which returns the programme to the Kaplan-Meier box.

Kaplan-Meier: Define Event for Status Var	iable 🗙
Value(s) Indicating Event Has Occurred           Image: Single value:         1	Continue
C Range of values:	Cancel
C List of values:	Help
Add	
Change	
Bemove	
,	

Click on the **Options** button on the Kaplan-Meier box, and select **Plots** > **Survival** (and possibly **Plots** > **Log Survival**) to get a graph. Now click on **OK** and SPSS will spend a while processing, and produce some numerical output in a separate Output window, ending with a survival curve. Note that the y-axis title should be edited to read 'Cumulative survival', instead of 'Cum survival'.

### Survival Function



Note for this example that the curve shows an initial, very sharp drop in survival, and then the slope appears to lessen with time.

However, this survival analysis is not very interesting, as it mixes together the 'Poison' and 'Non-treatment' groups. SPSS will separate these out, using the variable *Treatmen* as a factor, but it needs to be recoded as a number first. (This is just an old-fashioned feature of SPSS.) To recode, return to the SPSS Data Editor (the output is a separate SPSS Output Window).

File Edit	View	Data	Transform	Analyze	Graphs	Utilities	Window
□ □ □ □ □ 1 : band			Compute Random Count	e Number Se		=  <b>1</b>	
		bar	Categori	ize Variable		tre	atmen
1	47476	ì	Rank Ca	ses		oison	
2	47477	'	Automat	ic Recode.		oison	
3	47478	}	Create T	ïme Series		oison	
4	363F		Replace	Missing Val	lues	oison	
5	366F		Run Pen	ding Transf	forms	oison	
6	367E			ang mana	101.01	nison	

In the Data Editor, go to **Transform > Automatic Recode**:

Then select the variable *Treatmen*, and type a new variable name into the box toward the bottom right.

🗖 Automatic Recode		×
<ul> <li>▲&gt; band</li> <li>⊕ tx</li> <li>▲&gt; area</li> <li>⊕ on</li> </ul>	Variable -> New Name Irreatmen> ????????	OK <u>P</u> aste <u>R</u> eset
(₩) oft_last (₩) days (₩) Cdays* [cdays] (₩) Death? [death]	<u>N</u> ew Name Itreat Recode Starting from	Cancel Help

Click on **New Name** and then **OK**. This adds a new variable *treat* to the data with numbers instead of words for the two treatments. Now add the **new variable** as a **Factor**.

🗖 Kaplan-Meier		×
<ul> <li>▲ band</li> <li> ★ tx ▲ area ▲ treatmen ★ on ♦ off_last ♦ days </li> </ul>	Ime:         Cdays* [cdays]         Status:         death(1)         Define Event         Eactor:         Teat         Strata:         Strata:         Label Cases by:	OK <u>P</u> aste <u>R</u> eset Cancel Help
Compare Factor	Sa <u>v</u> e <u>O</u> ptions	

### Clicking on **OK** should lead to another lot of output:

Survival Analysis for CDAYS

Factor TREAT = Non-treatment

Time	Status		Cumulative Survival	Standard Error	Cumulative Events	Number Remaining
3		1	.9444	.0540	1	17
4		1	.8889	.0741	2	16
8		1			3	15
8		1	.7778	.0980	4	14
10		1	.7222	.1056	5	13
10		0			5	12
14		1	.6620	.1126	6	11
15		1	.6019	.1174	7	10
16		1	.5417	.1201	8	9
21		1	.4815	.1209	9	8
31		1	.4213	.1198	10	7
34		1	.3611	.1168	11	6
37		1	.3009	.1118	12	5
38		0			12	4
39		1	.2257	.1062	13	3
52		1	.1505	.0937	14	2
65		0			14	1
172		0			14	0
Number o	f Cases:	18	Censored:	4 (	22.22%) Events:	14
	Survival	Time	Standard Err	or 95% Co	nfidence Interval	_

Mean:	45		14	(	17,	72)
(Limited ) Median:	to 172 21	)	11	(	Ο,	42)
-						

Survival Analysis for CDAYS

Factor TREAT = Poison

Time	Status		Cumulative Survival	Standard Error	Cumulative Events	Number Remaining
4		1	.9744	.0253	1	38
12		1	.9487	.0353	2	37
14		1	.9231	.0427	3	36
15		1	.8974	.0486	4	35
16		1	.8718	.0535	5	34
16		0			5	33
17		1	.8454	.0581	6	32
17		0			6	31
18		1			7	30
18		1	.7908	.0659	8	29
21		0			8	28
21		0			8	27
22		1	.7615	.0697	9	26
22		0			9	25
23		1	.7311	.0732	10	24
26		0			10	23
31		1	.6993	.0766	11	22
32		0			11	21
34		0			11	20
35		1	.6643	.0804	12	19
47		0			12	18
51		1	.6274	.0840	13	17
53		1	.5905	.0868	14	16
68		0			14	15
70		0			14	14
72		0			14	13
73		0			14	12
76		0			14	11
80		1	.5368	.0940	15	10
96		0			15	9
106		1	.4772	.1007	16	8
112		1	.4175	.1043	17	7

Continued on next page

119 128 161	1 1 0	.3579 .2982	.1051 .1031	18 19 19	6 5 4
175 180	1 0	.2237	.1008	20 20 20	3
180	0			20	0
Number of -	Cases: 39	Censored: 1	.9 (48	.72%) Events	: 20
	Survival Time	Standard Error	95% Conf	idence Interva	1
Mean:	96	12	(	72, 120	)
Median:	106	34	(	39, 173	)
Survival	Analysis for CDA	YS			
		Total	Number Events	Number Censored	Percent Censored
TREAT TREAT	Non-treatme Poison	nt 18 39	14 20	4 19	22.22 48.72
Overall		57	34	23	40.35

This output has separate tables for the two treatment levels, and produces a graph with separate lines for them.



### Survival Functions

The curves look different. Note that each curve changes only when deaths occur, and that censored observations are individually marked. Again, the graph needs the y-axis label edited, and one of the lines needs to be changed to dashes, so that the lines are readily distinguished without colour (enabling black-and-white printing). This is achieved by double-clicking on the graph to open an editing window, carefully selecting just one of the lines, and changing the **Style** in the **Lines** tab of the **Properties** dialogue box.

### 5.3.2 Kaplan-Meier procedure using Excel

We now present Excel 2002 spreadsheets (Appendices 3 and 4) to handle the analysis of radio-telemetry data, designed for those who do not have access to standard statistical packages. While we have attempted to get things right, we cannot guarantee that these sheets will necessarily handle all datasets, nor that they will work in later versions of Excel. The actual spreadsheets used here are available by request from the senior author. Please acknowledge this paper if these Excel spreadsheets are used.

In this section, we will derive the Kaplan-Meier survival rates using the same example used for SPSS.

The key information for each individual is the same as for SPSS: the date of each event and whether there was a death or censorship. There may also be a covariate—in our example, Poison versus Non-treatment. In Excel, the process works best if there are no empty rows or columns in the main data area, and if any other information, such as derived totals, is separated from the actual data to be analysed for each individual by an empty row or column.

The first step is to select a cell in the main data area and create a pivot table: click on **Data > PivotTable and PivotChart Report**. (Note that the following instructions are for Excel 2002. Details may differ in different versions of Excel, but it will generally be possible to create the same table, providing the Pivot Table tool is available.)

×1 /	Microsoft Excel - WGNCR-37511 pracguide - tables.xls								
2	<u>E</u> ile <u>E</u> di	t <u>V</u> iev	∾ <u>I</u> n:	sert F <u>o</u> rmat	Tools	Data	a <u>W</u> indow <u>H</u> elp		
D	🖻 🖪	60	). Po	1 Kr + 🚷	Σ -	₽↓	<u>S</u> ort 8		
裕	ta ta i	2 6	1	0 10 2	₩¥Rep		<u>F</u> ilter		
	A2	-		<i>f</i> × 47476			Form		
	A	В	С	D	E	E Subtotals			
1	Band	Tx	Area	Treatmen	On		Validation		
2	47476	90	Rr	Poison	18-Sej				
3	47477	12	Rr	Poison	12-Sej		Table		
4	47478	29	Rr	Poison	13-No <sup>1</sup>				
-6	363F	96	Rr	Poison	21-0e		Text to Columns		
Б	366F	45	Rr	Poison	18-Oc		Constitute (		
1	367F	88	Rr	Poison	24-Oc		Consolidate		
8	371F	91	Rr	Poison	21-0c		Group and Outline		
9	372F	31	Р	Non-treatment	14-0c				
10	376F	17	Rr	Poison	18-Oc	17	PivotTable and PivotChart Report		
11	377F	96	Rr	Poison	14-Nov	_			
12	382F	27	Rr	Poison	6-Jai		Import External Data		
13	384F	94	Rr	Poison	17-Mar		Refrech Data		
14	385F	35	H	Non-treatment	2-May	÷	Kerresit Data		
15	386F	37	Р	Non-treatment	12-Sep	-97	27-Sep-97 15 15 1		

Go through the three steps of the wizard, checking at step 2 that **exactly** the rows and all the columns needed have been selected; extra columns do not matter. All the columns need meaningful headings. At step 3, the default is to put the pivot table on a new worksheet. It is generally a good idea to leave the data uncluttered on its own sheet. The skeleton of the pivot table will now be on the new sheet.



Now drag the time variable (*Cdays* in our example) to the left of the table where it says 'drop row fields here'. Drag the covariate factor if there is one (*Treatmen* in our example), then the event indicator (censorship or death; '*Death?*' here) to the top of the table to be the column fields, and ignore the page fields area. Also drop any variable with a complete set of values (i.e. having no empty cells) in the middle of the table. Preferably use one with character values, as this will default to giving the **count** of items that are required. Excel will automatically choose to **sum** a numeric variable, which will then have to be changed to a count. (In our case we used *Treatmen*.) It is important that the covariate (*Treatmen*) is to the left of the event indicator (*Death?*), which can be achieved by dragging the labels to put them in the correct order. This should produce the following table:

	A	В	С	D	E	F	G	Н	1
1			Drop	Page Fields Here					
2									
3	Count of Treatmen	Treatmen 👻	Death? 👻		_				
4		Non-treatment		Non-treatment Total	Poison		Poison Total	Grand Total	
5	Cdays* 🔻	0	1		ΪO	1	1		
6	3		1	1				1	
7	4		1	1		1	1	2	
8	8		2	DivotTable Field I	iek	- v		2	
9	10	1	1	THOUT ONLY THE OL				2	
10	12			Drag items to the P	ivotTable r	eport	1	1	
11	14		1	Rand			1	2	
12	15		1	Daiu		- 11	1	2	
13	16		1	TX		- 11	2	3	
14	17			Area		- 11	2	2	
15	18			Treatmen	1	- 11	2	2	
16	21		1	- E On		- 11	2	3	
17	22			E Off Last		- 11	1	1	
18	23					- 11	1	1	
19	26			Days		- 11	1	1	
20	31		1	Cdays*		- 11	1	2	
21	32			Death?		- 11	1	1	
22	34					- 11	1	1	
23	35						1	1	
24	37		1			_		1	
25	38	1		Add To Row A	rea	Ψ.		1	
26	30		1					1	

The Field List can now be closed, and the pivot table values can be used to derive the Kaplan-Meier survival rates.

We now describe in detail the formulae for creating the KM rates. With access to our spreadsheet, they can be copied over to apply to other data. However, it will be necessary to carefully check that the formulae refer to the correct cells. Our workings are shown in Appendix 3, Table A3.1. Note that in the formulae that follow, a cell reference (e.g. A7) generally must be typed in for cells in the pivot table, rather than selecting the cell to go into a formula, as Excel can create complicated references when a cell in a pivot table is selected.

#### The first group: 'Non-treatment'

- Column I: **Day** In the first column next to the table, create a copy of the time variable. In our example, we put =A6 in the cell **I6**, and copied it down the side of the table as far as, but not including, the Grand Total row. Next type 0 in the cell just above the first event day: in our example, **I5**. (This column will make things easier when we want to create a graph, as it is hard to select exactly the cells required from the pivot table itself.) Add a label, e.g. *Day*, immediately above this column, in cell **I4**. Type labels at the head of each column as shown in the table.
- Column J: At risk. Put the total number at risk (i.e. the sample size of the Nontreatment group) in the cell next to the first day of a record, in our case enter =D47 in J6. In the next cell down we take the value of the cell above, minus the value of total events, both censorships and deaths (i.e. 0s and 1s), from the preceding line. In our example, we put =J6-D6 in J7 and then copied this down. This gives the number at risk **before** any events on each day.
- Columns K and L: **Empirical death and survival rates**. The empirical death rate is simply the number of deaths divided by the number at risk. Thus, we divide the appropriate cell in column C by the cell in the same row in column J. For example, we typed =C6/J6 in K6, and then copied this down into all the appropriate cells. The empirical survival rate, which is essential, is 1 minus the death rate, so put =1-K6 in L6, and copy that down.
- Column M: **Kaplan-Meier (KM) product moment survival rate**. This is the information that we really want. Create it by taking the cumulative product of the survival rates to date. Type '1' in the row corresponding to time 0 (in **M5** in the example), to represent 100% survival rate initially. In **M6**, put =M5\*L6, and copy it down.

### The next group: 'Poison'

Now we have the key item we want for the first group. After putting useful headings on each column, as in our example, copy columns J to M into the next columns (N to Q) so the formulae can be slightly adjusted to give the KM survival rates for the second group (the references need to be corrected). The first cell in the new **At risk** column (N6) now must refer to the total in the second group (=G47 in our example). Adjust the next cell (N7) to refer to the cell immediately above, less the total number of losses (censorships and deaths) in this group (=N6-G6 here) and copy this down. The next column, empirical death rate, must be adjusted to ensure that it refers to the number of deaths in this group, divided by the number at risk. (Thus O6 has =F6/N6 in our example.)

There may be some entries '#*DIV/0!*' at the bottom of the table, if there are no longer any at risk in this group. In the example, this happens for the last couple of dates for the first group. It will pay to clear these problem cells (only) before graphing, as Excel will tend to interpret these as zero values. However, note that these formulae were required for the second group, and were used for copying.

Further useful calculations for standard errors and confidence intervals can follow, but first it is worth graphing these results. To graph the KM survival rates against time, put a short heading at the top of each group of KM rates. Select the time (column I), including the heading, and similarly the columns of KM survival rates. Now click on **Insert > Chart**, and select **XY (Scatter)**. This will give a graph of the survival curve like the one below:



This graph has been tidied up by adding labels; adjusting the vertical scale, the position of the legend and the size of each point; changing the symbols; and deleting the background and gridlines. It represents the survival rates adequately, but includes points at days where there are censorship events as well as points at each death. The SPSS graph shows the difference between the two types of events. Excel can also graphically show the differences between the types of events, if extra columns and the IF function are used, but we have not added this refinement.

The median duration of survival can be readily calculated by finding the time at which survival first drops below 0.5. Reading the values for the KM survival rate in the example, this is reached at 106 days for the Poison group, but at 21 days for the Non-treatment group.

The next step is to use Greenwood's formula (Klein & Moeschberger 1997: 84) to calculate the standard error and confidence intervals for the KM survival rate for each group. To do this for group 1, insert five columns after the KM rate for the first group. This and the following steps are shown in Appendix 3, Table A3.2 (Table A3.1 shows only the workings to this point).

- Columns N and O: **Two steps to Greenwood's formula**. The first step (to generate the values for column N) involves the following calculation for each group: divide the number of deaths by the product of the number at risk and the number at risk less the number of deaths. In the example, we put =C6/(J6\*(J6-C6)) in N6, and copied it down through both groups. The second step involves calculating the cumulative sum, by putting =SUM(N\$6:N6) in O6, and copying it down. Note that the \$ sign fixes a cell reference so that it does not change during copying.
- Column P: **Standard error**. This is simply the product of the KM survival rate and the square root of column O. Put =M6\*SQRT(O6) in **P6**, and copy it down through both groups. This gives a measure of the error in the KM rate. However, confidence intervals are often more useful.
- Columns Q and R: **95% confidence intervals**. Because survival rates should be between 0 and 1, it is best to use a different approach to the usual  $\pm$  1.96 × standard error (Klein & Moeschberger 1997: 97). Instead the formula for Q6 is =*M6^EXP(-1.96\*SQRT(O6)/LN(M6))* and for R6 is =*M6^EXP(1.96\*SQRT(O6)/LN(M6))*. Note that the only difference is the change of signs (- then +) after *EXP(.*

These columns can then be copied after the next group, and almost all the formulae will translate as needed. The only adjustment that should be required is a change in the first step of Greenwood's formula to ensure that it refers twice to the number of deaths in the correct group. In our example, the formula in **W6** should be =F6/(S6\*(S6-F6)), and this should be copied down.

That is all for the Kaplan-Meier procedure. Note that although SPSS does better charts, SPSS 12 does not give confidence intervals!

### 5.4 CHOICE OF METHOD

Every time a study animal, dead or alive, is found, the survivorship estimate will change. For day-to-day use, the Mayfield method is very easy to use and understand, the estimate can be easily recalculated (as in Appendix 1), and the method usually gives a reasonable picture of the survivorship rate. However, it is preferable that the more complicated calculations of the Kaplan-Meier procedure are used for reports or scientific papers that include survivorship data. Neither method is particularly accurate where there is a short total tracking time and / or a small number of deaths recorded. For example, there is a big difference between a survivorship rate of 3/20 = 0.15 and one of 4/20 = 0.20 caused by one more death of a study animal, especially if those data are then used to calculate life expectancy as 6.7 years versus 5.0 years, respectively. The number of deaths recorded is the most critical part of the calculations and, as a rough rule, aim to have either total tracking-years of the study being at least ten times the average life expectancy of the study animal, or the product of the number of deaths recorded and the number of tracking years exceeding 500 (e.g. 50 deaths in 10 tracking years' data, ten deaths in 50 years' tracking data or two deaths in 250 tracking years' data). However, remember that the fewer deaths recorded, the greater the change made by a chance event, or non-event; examining confidence intervals for the estimates gives a basis for evaluating the variability in the estimates due to chance.

# 5.5 HOW TO ESTIMATE SURVIVAL RATE TO A PARTICULAR AGE

There is often interest in calculating the survival of animals to a particular age (for example survival of kiwi chicks to 180 days old, at which time they seem to become reasonably safe from predation by stoats). The Kaplan-Meier procedure gives a survival rate as long as there are members of the group at risk. However, it can be subject to very large error when the sample size is small. For example, the estimate for Non-treatment survival in our kiwi chick sample is 0.1505 from 52 to 172 days, when the single chick left in the study was censored. The 95% confidence interval is (0.027, 0.370). For the Poison treatment, the survival rate estimate at 180 days is 0.2237, with the interval (0.067, 0.436). However, the Non-treatment estimate in particular is based on very few data.

If constant survival rate is assumed, the Mayfield method can be used to estimate a survival rate at any point. This assumption can be checked by looking at a Kaplan-Meier graph with the survival axis on a log scale. The SPSS option to get the log survival curve is described above. In Excel, double click or right click on the vertical (y) axis to bring up the **Format Axis** dialogue box, choose the **Scale** tab, and select **logarithmic scale**. To assess whether the points are reasonably consistent with a straight line, look mainly at the points that correspond to actual event, rather than censorship.

The constant survival estimate of survival rate to time *t* is:

 $\exp(\frac{-t}{L})$ where *L* is the simple estimator of life expectancy described earlier (i.e. total time exposed *T*, divided by number of deaths observed *d*). A confidence interval can be calculated for this also, and the results of doing this for the kiwi chick data are shown in Appendix 4 with estimates at 180 days. The data in Appendix 4 are derived from Appendix 2 using a simple pivot table. Note that the estimates are similar to, but not the same as, the Kaplan-Meier estimates, as a different model is being used. In particular, it appears to give a more realistic estimate of survival to 180 days for Non-treatment. It is important to note that the validity of this confidence interval is heavily dependent on the assumption of constant survival.

The sheet used to create Appendix 4 can be modified for other data by entering the appropriate values where there are numbers in bold: the total time exposed, the number of deaths, and the point at which the estimate is desired. The formula used here for the confidence interval is:

$$\left\{ \exp\left[\frac{-t}{2T}\chi_{2d,\frac{\alpha}{2}}^{2}\right], \exp\left[\frac{-t}{2T}\chi_{2d,\left(1-\frac{\alpha}{2}\right)}^{2}\right] \right\}$$

following the same notation as given above, in section 5.2 (Lawless 1982). In Excel, the point estimate is given by the formula =EXP(-t/L), and the 95% confidence intervals are given by =EXP(-t/(2\*T)\*CHIINV(0.025,2\*d)) and =EXP(-t/(2\*T)\*CHIINV(0.975,2\*d)). When typing these formulae into Excel, the appropriate cell references must be placed where the references *L*, *d*, *t* and *T* are given above.

# 6. Comparison between two or more groups

As an extension of the Kaplan-Meier procedure, it is possible to compare the survivorship of animals in two or more different groups, e.g. males versus females, or animals living under a number of different management regimes. The most appropriate statistic to use is the nonparametric Mantel-Haenszel statistic, which is a log-rank test whose distribution approximates a  $\chi^2$  distribution with 1 degree of freedom for two groups, or (G - 1) degrees of freedom if there are G groups. The statistic is computed by combining the two (or more) samples to be compared. It is then determined whether the times when deaths were recorded in the two groups are sufficiently different from one another (given the number of animals at risk in each group at each age that an animal died).

To carry out this test in SPSS, continue the previous SPSS analysis by simply clicking on the **Compare Factor** button on the Kaplan-Meier dialogue box, and select **Log rank > Continue** to exit that box, and **OK** to run the survival analysis again.

Kaplan-Meier: Compare Factor Levels	×
Test Statistics         ✓ Log rank       ☐ Breslow       ☐ Tarone-Ware         ✓ Linear trend for factor levels         ④ Pooled over strata       ○ Pairwise over strata         ○ For each stratum       ○ Pairwise for each stratum	Continue Cancel Help

SPSS has a habit of hiding some of the text output at the bottom of the output. If this happens, try selecting and resizing the box with the output text in it, or run the analysis again with the other output options turned off, so that only the results of the log-rank test are provided. In our previous example, this will give the following SPSS output:

```
Test Statistics for Equality of Survival Distributions for TREAT
Statistic df Significance
Log Rank 8.60 1 .0034
```

Given that a probability, *P*, of 0.0034 is well below the accepted statistical threshold of 0.05, we conclude that kiwi chicks in the areas treated with brodifacoum poison survived significantly better than in unpoisoned blocks nearby. This was probably because stoats (*Mustela erminea*) and cats (*Felis catus*), the main predators of young kiwi, were killed by secondary poisoning after eating dead or dying rats or possums, and this clearly outweighed any risk from accidental poisoning of the kiwi chicks themselves.

The steps in using Excel to compare two or more groups are described below and refer to the spreadsheet in Appendix 5. The data presented is the same chick survival data used earlier (Appendix 2).

We start by using the same pivot table as for the Kaplan-Meier procedure. Either follow the instructions above to create an identical pivot table or copy the pivot table shown in Table A3.1 to a new sheet and remove all the workings. (It would be possible to add this material to the KM table, but it could become difficult to follow.)

- Column I: **n1j**. The number at risk in the first group (Non-treatment) at that time (it includes the animal that died and any animals censored at exactly that time). This is calculated from the total in the group, less earlier deaths and censorship. The Excel formula in the first cell in the example is =D\$47-SUM(D\$5:D5), and should be copied down. The formulae for the first cell for columns J-Q is given below in italics, and should be copied down the sheet.
- Column J: **n2j**. The number of animals at risk in the second group (=G\$47-SUM(G\$5)).
- Column K: **nj**. The combined total number of chicks at risk (=I6+I6).
- Column L: **dj**. The total number of deaths at that particular tracking time (the sum of **d1j**, the number of deaths in group 1 from the pivot table column C, and **d2j**, the number of deaths in group 2 from the pivot table column F (=C6+F6).
- Column M: e1j. The expected number of deaths in groups 1 at that particular time if survival was the same in the two groups (=L6\*I6/K6).
- Column N: e2j. The expected number of deaths in group 2 (=L6\*J6/K6).
- Column O: d1j e1j. The difference between observed deaths and expected deaths in group 1 (=*C6-M6*).
- Column P: d2j e2j. The difference between observed deaths and expected deaths in group 2 (=*F6-N6*).
- Column Q: The **estimated variance** of the differences (=I6\*(K6-I6)\*(K6-I6)\*(K6-I6))).

(Note that columns N and P are unnecessary for this two-sample test because they are the complement of columns M and O. They are shown here because they would be needed if the number of groups was greater than 2.)

Next, sum the differences and square this sum to get the test statistic: add up values in column O or column P (= ± 6.83 in the kiwi example), square the answer (= 46.64) and divide it by the sum of variances (add up values in column O (= 5.421)); this gives a test statistic of 8.60. This figure can then be compared with the percentile values in statistical tables of the  $\chi^2$  distribution with 1 degree of freedom. Excel calculates these for us, using the function *CHIDISTO*, with the test statistic as the first argument and the degrees of freedom as the second argument. From this kiwi example, we concluded that chick survival was significantly better in the poisoned areas than in nearby unpoisoned areas (*P* = 0.0034), as previously reported by Robertson et al. (1999).

# 7. Further topics in survival data analysis

This guide provides some simple robust tools for analysing survival data. The Kaplan-Meier survival rate and log-rank test described herein do not include specific assumptions about the distribution of survival times. In addition to these tools, there are other very well-developed tools available for survival data analysis, which can extract further information from various sorts of data. One approach involves making specific assumptions about the way survival times are distributed—leading to parametric models, such as the exponential model used above, and the more flexible Weibull model. Another very important direction involves using the Cox proportional hazard model, which allows the inclusion of various covariates in a semi-parametric model. As more survival studies are designed, implemented and their results analysed, some of these more advanced approaches may be needed, but the tools covered in this guide should provide a good starting point.

## 8. Availability of Excel files

Copies of the Excel files used in Appendices 1-5 are available by request from the senior author. The files can be saved and new data substituted for the old in the files—taking great care not to write over formulae—and after some adjustments to data references the calculations will be done automatically. Alternatively, the data could be copied into Excel from an electronic (pdf) version of this manuscript. Use the text import wizard (Data > Text to columns... using Space as the delimiter) to recreate the data tables in Excel and follow the instructions given. Similarly, the more complex formulae could be clipped from the pdf into Excel.

Radio-tracking has become a very powerful tool for determining the survivorship of wild animals. It is free from many of the assumptions inherent in other methods of calculating survival using capture-recapture techniques. Researchers must, however, be ever vigilant, because catching wild animals, attaching transmitters to them and regularly radio-tracking them (with its various levels of disturbance) may affect the survival chances of the study animal. It is important to keep up with improvements in transmitter technology, packaging and attachment methods, and to use mortality transmitters wherever possible. If the chances of mortality are increased through an animal wearing a transmitter, survival estimates will be conservative, whereas with other methods biases can lead to either conservative or inflated survival estimates.

The main problem with radio-tracking studies is obtaining a sufficiently large sample of animals and—most importantly for long-lived species—getting a sufficient number of recorded deaths to make the estimates reliable. The methods described above require considerable time in the field to obtain good survivorship estimates; however, the aim of most radio-telemetry studies is for more than just collection of survivorship information, to 'kill two (or more) birds with one stone'!

The tests described here can also be used in some other situations where animals are marked in other ways and then recaptured/resighted later. However, be aware that with some methods the assumptions can be seriously violated, e.g. birds often avoid recapture in mist-nets, and this can create serious problems with capture-recapture analysis. It will often be best to use specialist software, which is now available for this sort of data. The statistical methods presented here also seem to be appropriate for studies of plant survival, permitting, for instance, comparison of the survival of tagged or counted plants in one plot or quadrat with another (e.g. grazed versus ungrazed) at various (regular or irregular) intervals.

## 10. Acknowledgements

The work that led to this paper was carried out as part of Bank of New Zealand Kiwi Recovery, which is a cooperative project between the Department of Conservation and the Bank of New Zealand (DOC Science Investigation No. 1218). Pat Miller did most of the day-to-day radio-tracking of brown kiwi in central Northland to obtain the data presented as the examples here. Robert Gentleman (Harvard School of Public Health) provided initial statistical advice. Pat Coope (Statistics New Zealand), Ralph Powlesland (DOC), Sue Hallas (DOC) and two anonymous referees provided useful testing of the draft programmes and / or comments on the manuscript.

### 11. References

- Bunck, C.M.; Chen, C.-L.; Pollock, K.H. 1995: Robustness of survival estimates from radiotelemetry studies with uncertain relocation of individuals. *Journal of Wildlife Management* 59: 790–794.
- Clobert, J.; Lebreton, J.-D. 1991: Estimation of demographic parameters in bird population dynamics. Pp. 75-104 in Perrins, C.M.; Lebreton, J.-D.; Hirons, G. (Eds): Bird population studies: their relevance to conservation and management. Oxford University Press, Oxford.
- Clout, M.N.; Karl, B.J.; Pierce, R.J.; Robertson, H.A. 1995: Breeding and survival of New Zealand pigeons *Hemiphaga novaeseelandiae*. *Ibis 137*: 264-271.
- Heisey, D.M.; Fuller, T.K. 1985: Evaluation of survival and cause-specific mortality rates using radio-telemetry data. *Journal of Wildlife Management 49*: 668–674.
- Klein, J.P.; Moeschberger, M.L. 1997: Survival analysis: techniques for censored and truncated data. Springer-Verlag, New York. 502 p.
- Lawless, J.F. 1982: Statistical models and methods for lifetime data. Wiley, New York. 580 p.
- Lebreton, J.-D.; Burnham, K.P.; Clobert, J.; Anderson, D.R. 1992: Modelling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological Monographs* 62: 67-118.
- Mayfield, H.F. 1961: Nesting success calculated from exposure. Wilson Bulletin 73: 255-261.
- Mayfield, H.F. 1975: Suggestions for calculating nest success. Wilson Bulletin 87: 456-466.
- Miller, H.W.; Johnson, D.H. 1978: Interpreting results of nesting studies. *Journal of Wildlife Management 42*: 471-476.
- Pollock, K.H.; Winterstein, S.R.; Bunck, C.M.; Curtis, P.D. 1989a: Survival analysis in telemetry studies: the staggered entry design. *Journal of Wildlife Management* 53: 7-15.
- Pollock, K.H.; Winterstein, S.R.; Conroy, M.J. 1989b: Estimation and analysis of survival distributions for radio-tagged animals. *Biometrics* 45: 99-109.
- Robertson, H.A.; Colbourne, R.M.; Graham, P.J.; Miller P.J.; Pierce, R.J. 1999: Survival of brown kiwi exposed to brodifacoum poison in Northland, New Zealand. *New Zealand Journal of Ecology 23*: 225–231.
- Trent, T.T.; Rongstad, O.J. 1974: Home range and survival of cottontail rabbits in southwestern Wisconsin. *Journal of Wildlife Management* 38: 459-472.