

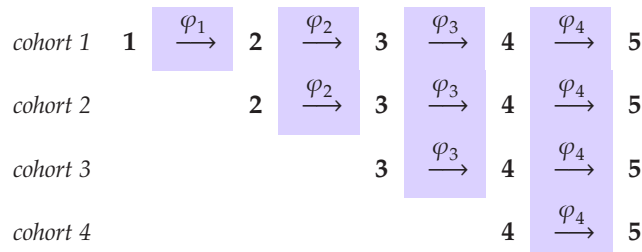
CHAPTER 7

‘Age’ and cohort models...

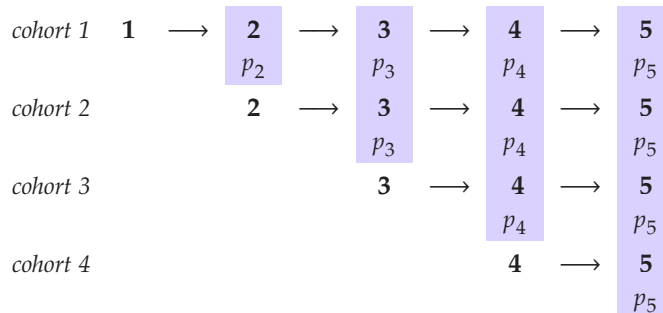
To this point we have assumed that our ‘underlying’ model has been the CJS time-dependent live encounter model. However, there are many instances when the ‘typical assumptions’ of the CJS model are not met. Specifically, the assumptions concerning independence of fate and identity of rates among individuals. Recall that the CJS model generally assumes that ‘all individuals, whatever their age or capture history, should have the same probabilities of capture and survival’, over a given interval.

We do not expect this to be generally true. In this chapter, we consider models which account for two common sources of differences among individuals: one that changes over the course of an individuals life (**age**), and one that does not (**cohort**). In both cases, we anticipate that differences in the age or cohort of an individual may influence its survival or encounter probability. For example, we might have strong reason to suspect that the survival of a young individual may generally be lower than the survival of an older individual.

We begin by reviewing the standard CJS model structure for a simple live encounter study. Recall the basic structure of the CJS time-dependent model – shown below for apparent survival, φ_i :



and for recapture (encounter) probability, p_i



Under the 'standard CJS assumptions', we assume that survival and recapture potentially vary only as a function of the time interval (as indicated by the shaded columns in the preceding figures). However, clearly this might not always be the case – both the time interval *relative* to when the individual was marked, and the time at which the individual was marked, may also be important. We start with the former – 'time since marking' (TSM).

7.1. 'Age' models

It will be trivially obvious to most biologists that individuals of different age classes (or developmental stages) are likely to differ in the probability of surviving to the next age or stage. In fact, life history theory is to a large degree focussed on analysis of such differences. The reasons for this are well-established. Organisms of a given age (for simplicity, we will refer only to age transitions – the logic applies reasonably well to simple stage-structured systems as well) may be more or less vulnerable to sources of mortality than are individuals of different age(s). The reasons for these differences might reflect differences in size, behavior, or physiological maturation. It is probably safe to say that there have been as many papers in the ecological and evolutionary literature devoted to 'age-dependence' of one trait or another as any other subject. So, we need to be able to address the question: are there age-specific differences in survival or encounter probabilities? To do that, we need to consider the construction of 'age' models (the reason for referring to 'age' parenthetically will become clear later in the chapter).

To begin, some definitions. Clearly, the ageing process begins when individuals are born. All individuals born in a given breeding season can be said to belong to the same birth cohort. A *cohort* is simply some criterion by which individuals are grouped together (birth year, in this case). Within a birth cohort, age and time are logically synonymous. In fact, age, time (e.g., year) and cohort are related by the following expression:

$$\text{age} = \text{current year} - \text{birth cohort}.$$

Consider the first row (i.e., first cohort) extracted from the triangular matrix (PIM) of a simple CJS time-dependence survival model (say, those shown on the previous page):

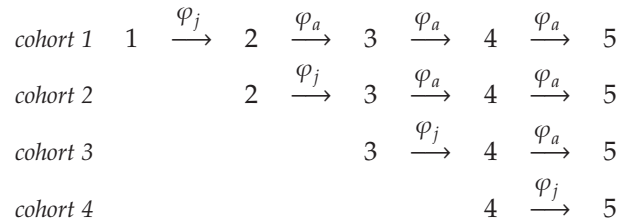
$$\text{cohort 1} \quad 1 \quad \boxed{\varphi_1 \rightarrow} \quad 2 \quad \boxed{\varphi_2 \rightarrow} \quad 3 \quad \boxed{\varphi_3 \rightarrow} \quad 4 \quad \boxed{\varphi_4 \rightarrow} \quad 5$$

This first row also corresponds to the first release cohort. Let's assume that this is also a 'birth' cohort. All individuals that are newly marked and released at occasion 1 were marked as newborns. Let's also assume (for simplicity) that the sampling occasions correspond to 'years'. Thus, individuals marked at occasion 1 (0 years of age) are, if they survive, 1 year of age at occasion 2, 2 years of age at occasion 3, and so forth.

Now, the parameters shown in this table ($\varphi_1 \rightarrow \varphi_4$) were originally written to show simple time-dependence. But, since individuals also age through time, we cannot differentiate between age-specific differences in survival, and simple time-specific differences. They are entirely confounded.

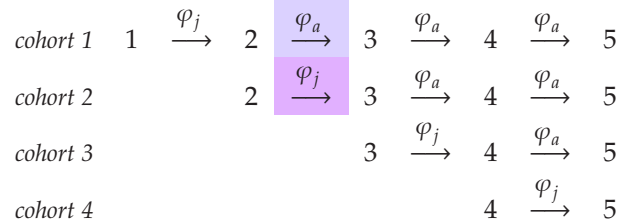
How do we then separate age and time effects? Clearly, this cannot be accomplished using a single cohort alone. However, what happens if we use multiple birth-cohorts? To examine this situation, let's make the following assumptions about some arbitrary population. First, let's assume that all individuals in each release cohort are marked as newborns. Let's also assume that survival between age 0 and age 1 year is different from survival after age 1 year. For simplicity, let's call the survival from 0 \rightarrow 1 'juvenile' survival, and survival from any age x (where $x \geq 1$ year) to $x + 1$ 'adult' survival.

If you think about it, this is not at all an uncommon situation. However, let's make it even simpler. Let's assume for the time being that juvenile survival is constant among cohorts, and that adult survival is constant both within cohorts and among cohorts. What would the parameter structure of this model look like? Let's use the '*j*' subscript for juvenile survival, and the '*a*' subscript for adult survival.



Now, let's look at this table and see how it makes sense. The best starting point is to look at each cohort separately. Within a release (birth) cohort, we see that between the first occasion (in the cohort), and the second occasion after marking, individuals survive with probability φ_j . However, from the second occasion after marking onwards (within a cohort), they survive with probability φ_a .

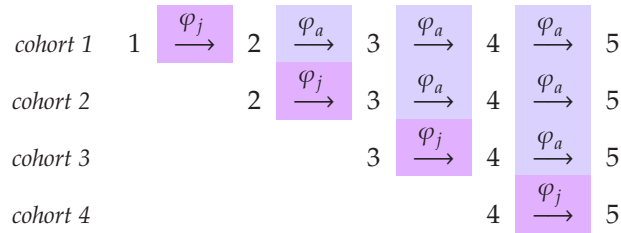
Now, as we discussed previously, within a cohort we cannot differentiate between 'time' and 'age'. However, note that we can test whether a model with this structure differs from one where (say) survival is constant (no age or time effect). But what we're really after is the affect of 'age' on survival, as separate from 'time' effects. Here is where multiple cohorts come in. By contrasting parameter estimates *among* cohorts, but *within* a time interval, we can differentiate age and time effects. For example, concentrate on the interval between occasion 2 and occasion 3 (shaded column/interval – below).



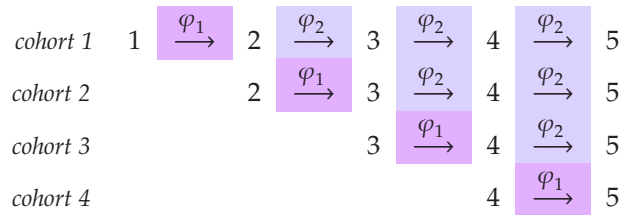
In terms of time, both cohorts 1 and 2 are experiencing the same 'temporal effect'. In other words, all individuals, whether they were newly marked at occasion 1 or occasion 2, are experiencing the aspects or characteristics of the 'environment' causing mortality during this interval. However, in cohort 1, the individuals at occasion 2 are 1 year of age ('adult'), while those from cohort 2 are newborn ('juvenile'). Thus, for cohort 1 individuals, they will survive from occasion 2 \rightarrow 3 with probability φ_a . In contrast, for cohort 2 individuals, they are surviving from occasion 2 \rightarrow 3 with probability φ_j .

If there were no age-specific differences in survival, then the ratio of survival of cohort 1 individuals over this interval would be the same as the survival of cohort 2 individuals over this interval. In other words, the ratio of φ_a/φ_j would equal 1. Again, it is the contrast *among* cohorts (i.e., rows), but *within* columns (i.e., intervals between occasions) that allows us to test for age differences in survival. This is a very important concept to grasp, so it is critical that you spend the time now to make sure you do.

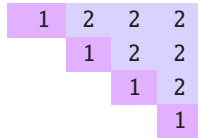
Let's start to formalize our notation a bit. First, consider the question 'what is the time axis of our model?'. The answer depends largely on which age class you are focussed on. For example, consider again our model with constant juvenile and constant adult survival, we essentially have 2 parameters. The differently shaded areas of our model (top of the next page) show the juvenile (along the diagonal) and adult age classes (above the diagonal), respectively.



The juvenile age class in this example spans one time interval (e.g., one year). The adult age class (above the diagonal) spans $(n - 1)$ years, where n is the number of occasions, and '1' is the duration of the juvenile age class. If we re-write the parameter matrix using numbers for subscripts instead of letters (let '1' = 'j', and '2' = 'a'), then our model with constant juvenile and adult survival would be

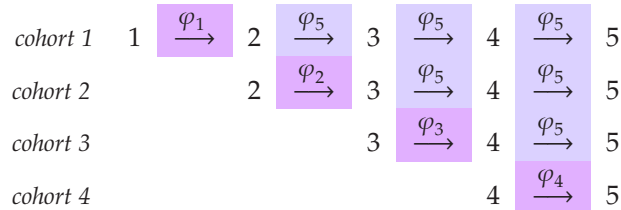


Now, if we simplify this and re-write the parameter structure the way that **MARK** interprets it (using only the subscripts), this 2-age-class model (juvenile, adult) is written as:

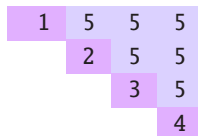


OK, now let's expand our model somewhat, adding some more 'flexibility'. Suppose for example that juvenile survival varies over time, but that adult survival is constant through time.

If we now add time-dependence to the juvenile survival probability, but leave adult survival constant, the model structure would now look like:

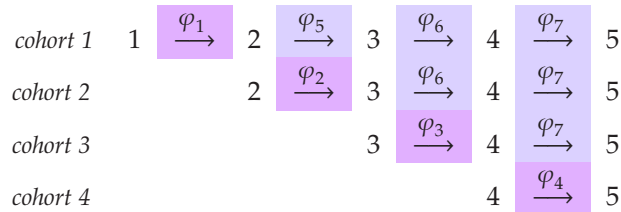


In the **MARK** parameter (PIM) format, this would reduce to:

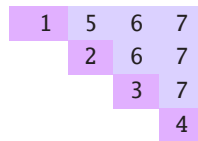


Let's extend it one more step: let's add time dependence to adult survival as well. This particular model is important since 2 age-class models, with a juvenile age-class spanning one year, and a single adult age class, with time-dependence in both, are very commonly seen in analysis of mark-recapture data from wild populations.

Here is what the 2 age-class model with time-dependence in both age classes would look like:

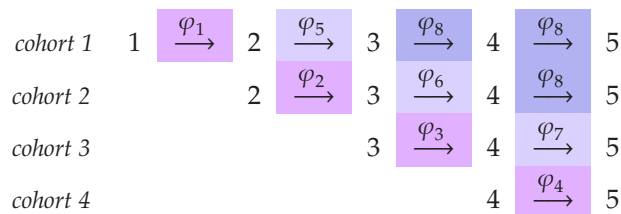


In the **MARK** parameter (PIM) format, this would reduce to:

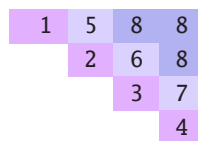


As a final test, to see if you really understand the structure of these models, consider the following situation. We have a 5 occasion study of a long-lived organism with indeterminate growth, and we believe that survival may be age-dependent. We decide to model survival in the following way. 3 age classes, the first age class spanning 1 year, the second age class spanning 1 year, and the final age class spanning all remaining years. In other words, a single-year duration 'juvenile' phase, a 1-year duration 'sub-adult' phase, and the final 'adult' phase. Juvenile and sub-adult survival are time-dependent, but adult survival is constant over time.

Here is the structure for this model:



which would reduce to:



With a bit of thought, you should be able to see how this model was constructed (look carefully at the ordering of the parameter subscripts). If not, go back through the preceding few pages, and try again. Age models are very important.

If you 'get the basic idea', then let's proceed to analysis of some simple 2 age-class models. We will examine how to modify the standard CJS PIMs for different types of age models – modifying the PIMs will let you construct an age (or cohort) model of arbitrary design. We will use the sample data set **age.inp**. We 'suspect' that there are 2 age-classes in these data. We want to 'confirm' our suspicion by using **MARK** to test the fit of various 2 age-class models versus the standard time-dependent CJS model with no age effects. There is only one group in the data set, and 7 sampling occasions.

Our candidate models (i.e., the list of those models that we believe, based on our biological expectations, might be appropriate to these data) are tabulated below. Note the subscripting conventions – with more 'structure' in our models (i.e., time, age, cohort), the subscripting can get a bit tricky to handle in any intelligible fashion. We've adopted a convention where the age classes for a given parameter are indicated by a 'frontslash' (/). For example, $\varphi_{t/.}$ would indicate 2 age classes for apparent survival, with time variation for the first age class, and time invariance ('dot') for the second age class.

<i>model</i>	<i>description</i>
$\varphi_t p_t$	standard CJS model - no 'age structure' – time dependence in both survival and recapture
$\varphi_{./} p_t$	2 age-classes for survival – both age classes constant (./.) through time. Time-dependent recapture.
$\varphi_{t/t} p_t$	2 age-classes for survival – both age classes time-dependent (t/t) through time. Time-dependent recapture.
$\varphi_{t/.} p_t$	2 age-classes for survival – juvenile (first) age classes time-dependent, adult age class constant over time (t/.) through time. Time-dependent recapture.

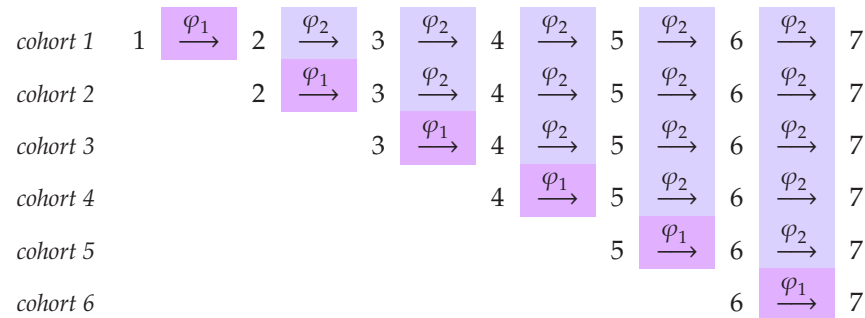
Clearly, we are focussing only on the survival side of things with these 4 models. However, do not get the idea that age-effects are only relevant to survival. They're not, and are equally as likely to show up in recapture probabilities as well. As a general 'rule' – do not overlook modeling the recapture side of things with as much interest and care as you do the survival side. We often tend to focus on the 'finality' of a survival analysis (since death has rather obvious fitness consequences), but why an individual isn't seen on a given occasion can be of equal interest.

By now, you should be able to start **MARK** and run the standard CJS model almost by reflex – go ahead and do so, and add the results to the browser. The CJS model makes a reasonable 'null' to test against – it is well-parameterized, but has no 'age structure'. Once the CJS model run is complete, we'll proceed and run the other 3 candidate models, in order, starting with $\{\varphi_{./} p_t\}$ – 2 age-classes for survival, constant over time for both ages, and time-dependent recapture probability.

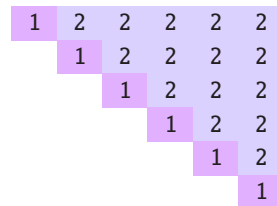
How do we construct and fit this model in **MARK**? By now, you should recognize that virtually *all* models in **MARK** are set by manipulating the PIMs and/or the design matrix, either alone or in combination (if you don't realize this, for shame – and go back and re-read Chapter 6). In this case, since all we're doing is changing the way in which the parameters are indexed (to reflect the age-structure), we modify the PIMs. So, bring up the survival and recapture PIMs.

Obviously, the recapture PIM will 'look' like the standard CJS PIM we've seen many times already – full time-dependence (this is what **MARK** defaults to). But what about the survival PIM? How do we modify to reflect the model structure $\{\varphi_{./} p_t\}$? Actually, for this particular model, it's pretty easy. **MARK** will let you construct this model using one of its built-in menu options. As a first step, you should be able to write out what the PIM should look like before you construct it.

In this case, with 7 occasions, the PIM should reflect the following structure:

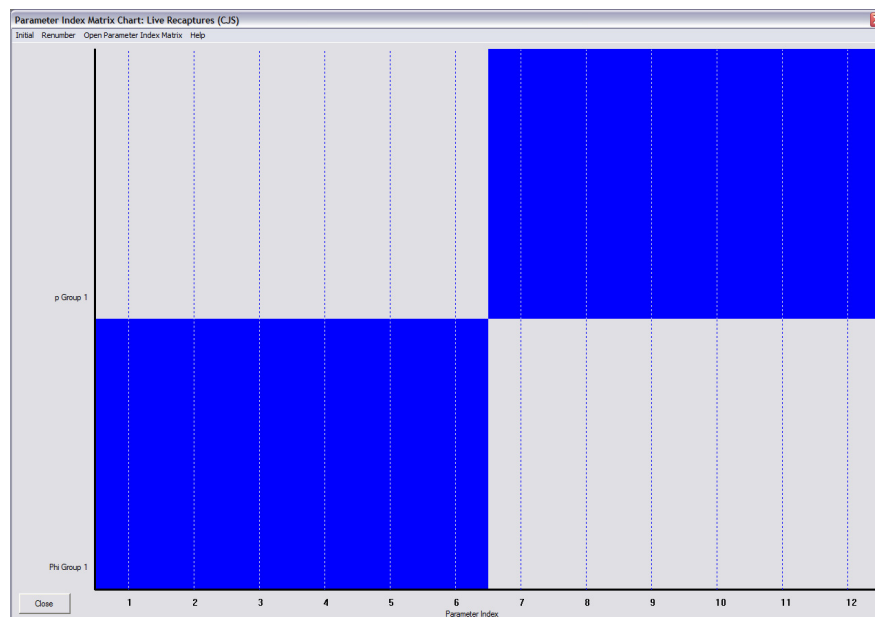


which corresponds to the following PIM structure...

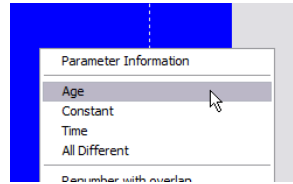


Make sure you understand the connection. Pay particular attention to the parameter subscripting (which of course leads to the parameter indexing you need to keep track of when constructing the PIMs in **MARK**, and reading the output).

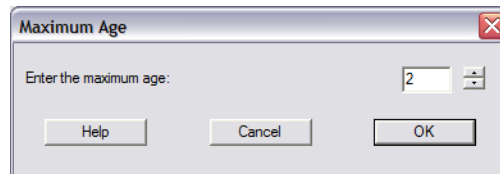
Again, while it would be relatively straightforward to modify the survival PIM to reflect this structure by manually editing each cell, **MARK** makes it much easier for you than that. As we've seen in earlier chapters, one of the handy features of **MARK** is the ability to modify the PIMs through the '**Parameter Index Chart**' (PIM chart). Open up the PIM chart, which initially reflects the CJS time-dependent model.



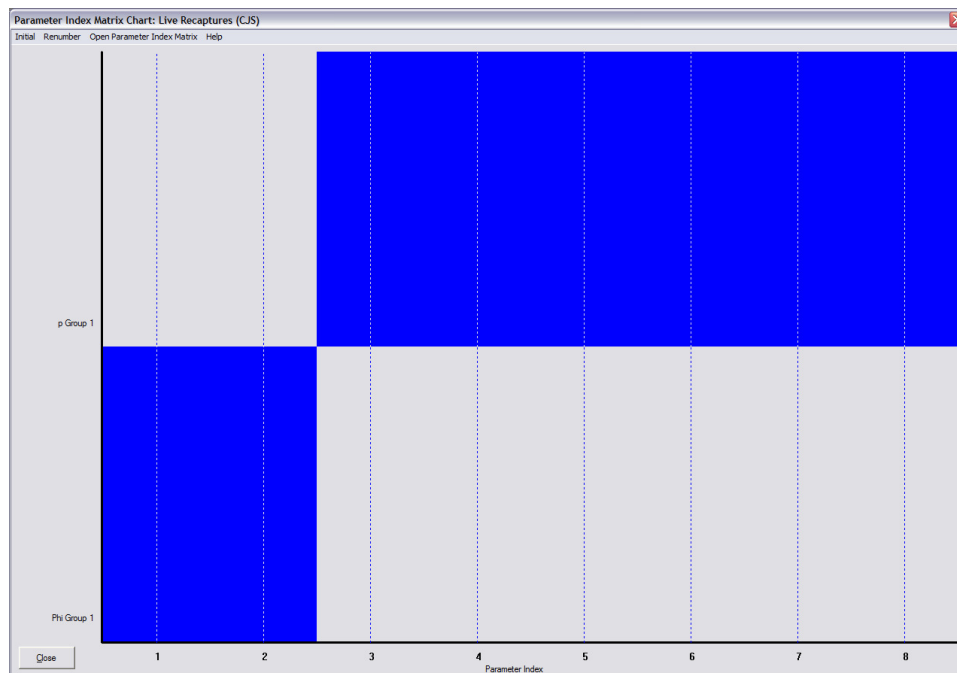
Moving the cursor over the survival 'blue box' in the PIM chart (the one in the lower left-hand corner), right-click the mouse. This will spawn a menu allowing you to select from a variety of 'built-in' parameter structures:



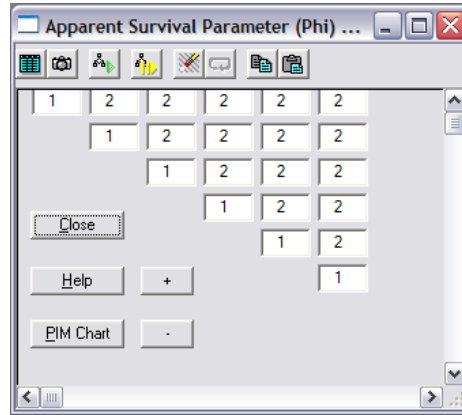
Obviously, the '**Age**' model is the one we're interested in here, so go ahead and select '**Age**' from this menu. You will then be presented with a small dialogue window asking you to define the age of the oldest age class. Now, this is potentially confusing. If each age class spans one year, this is the same as asking 'how many age classes do you want'? In our analysis, we want 2 age classes: one for 'juveniles' or 'young' (i.e., the first year after marking as offspring), and 'adult'.



Once you've clicked '**OK**', the PIM chart will reflect this change. To eliminate any 'gap' between the two parameters in the PIM chart, you can simply left-click and drag the box corresponding to recaptures over to the left. Ultimately, the PIM chart should look like:



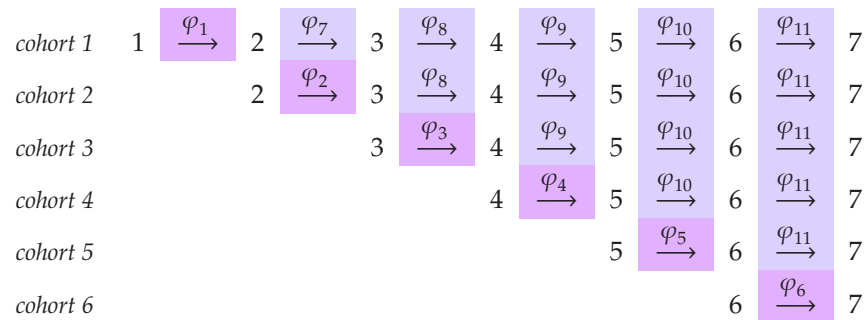
Now, in some senses, the PIM chart doesn't give you the best indication of the actual PIM structure. To look at it (and to confirm the PIM is structured correctly), right-click again over the survival 'box' in the PIM chart, and open the corresponding PIM window (one of the menu options). Then, close the PIM chart and have a look at the survival PIM:



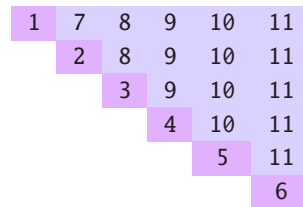
As you can see, it looks exactly as we wanted it to (check back earlier in this chapter if you don't remember the details for this model). Go ahead and run this model – we'll use the default sin link, and name the model 'phi(. /.)p(t)'. Add the results to the browser.

The next model is $\{\varphi_{t/t} p_t\}$ – 2 age-classes, but this time with full time-dependence within each age-class. Time-dependent recapture.

Here is the structure of this model (and the corresponding PIM):



which corresponds to...

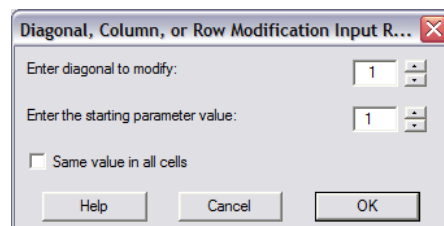


How do you get the PIM to look like this in **MARK**? There is no simple menu option for this model (the menu we used for the previous model only allows you to build age models which are constant over time in each age class).

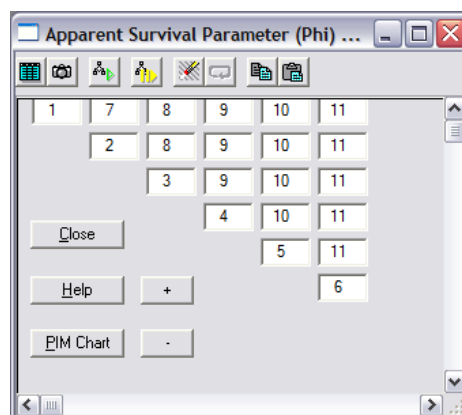
However, there is a relatively simple way to construct the PIM. First, you need to pay attention to the parameter indexing. Note that the largest value on the diagonal representing the first age-class is '6', and the first adult parameter value is (therefore) '7'.

So, here's what you do:

1. make the first cell value '6', instead of '1'
2. pull down the '**Initial**' menu, and select '**Time**'. This will create a PIM that is a standard CJS time-dependent PIM, but one that starts with '6'.
3. next, go back to the first cell, and click it again. Pull down the '**Initial**' menu, and select '**Diagonal**'. When you do, **MARK** will present you with a window asking you which diagonal you want to modify – since you've clicked in the first cell, it will default to the first diagonal, but you can always change this. It will also ask you what you want the starting parameter value to be. Change it to '1', and click '**OK**'.

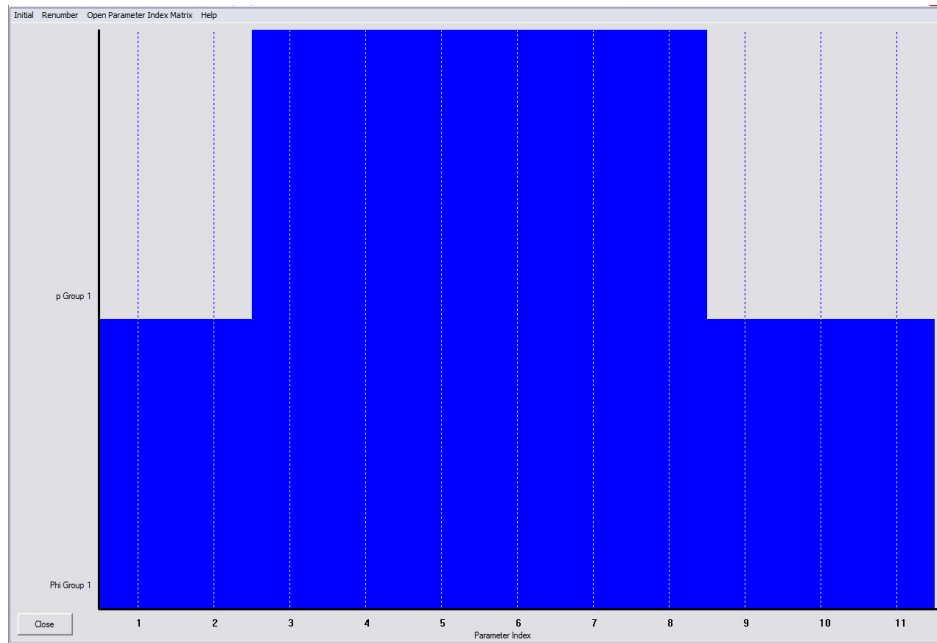


Now, have a look at the PIM – it should look exactly like we want it to – time-dependent indexing along the first diagonal from 1 → 6, and then time-dependence for the second 'adult' age class, from 7 → 11.



What we just did makes use of the fact that this model is basically a modification of the CJS model – in fact, the classic time-dependent structure within cohort is maintained for the second age-class, so all you do is figure out what the first parameter value would need to be for this age class, and go from there. With a bit of thought, we think you'll get the hang of it.

Now, before you run this model – have a look at the relative indexing of the survival and recapture PIMs. You can do this easily by using the '**Parameter Index Chart**', from the '**PIM**' menu:



What do we see? We see that the parameter indexing is 'overlapped' between the 2 parameters – this is **not** good, and is not something that **MARK** will 'fix behind the scenes'. As such, you **must** correct this problem yourself. The easiest way to do this is by modifying the 'overlap' via the PIM chart itself. You can either (i) drag the recapture 'box' to the right, or (ii) use the '**Renumber without overlap**' option. Simply right-click anywhere in the PIM chart, and select '**Renumber no overlap**'. Once you've corrected the indexing, go ahead and run the model – name it ' $\Phi(t/\tau), p(t)$ ', and add the results to the browser.

The final model is $\{\varphi_{t/}, p_t\}$ – 2 age-classes, with time-dependence in the first age-class, but constant survival in the second age-class. This model is very common for analysis of individuals marked as young – temporal variation in the first age class, but no variation among 'adults'. The idea, of course, is that individuals may be more susceptible to annual variation in conditions affecting survival in their first year than they are if they survive to adulthood.

Here is the parameter structure for this model:


cohort 1	1	$\varphi_1 \rightarrow$	2	$\varphi_7 \rightarrow$	3	$\varphi_7 \rightarrow$	4	$\varphi_7 \rightarrow$	5	$\varphi_7 \rightarrow$	6	$\varphi_7 \rightarrow$	7
cohort 2			2	$\varphi_2 \rightarrow$	3	$\varphi_7 \rightarrow$	4	$\varphi_7 \rightarrow$	5	$\varphi_7 \rightarrow$	6	$\varphi_7 \rightarrow$	7
cohort 3					3	$\varphi_3 \rightarrow$	4	$\varphi_7 \rightarrow$	5	$\varphi_7 \rightarrow$	6	$\varphi_7 \rightarrow$	7
cohort 4							4	$\varphi_4 \rightarrow$	5	$\varphi_7 \rightarrow$	6	$\varphi_7 \rightarrow$	7
cohort 5									5	$\varphi_5 \rightarrow$	6	$\varphi_7 \rightarrow$	7
cohort 6											6	$\varphi_6 \rightarrow$	7

which corresponds to...

1	7	7	7	7	7
	2	7	7	7	7
		3	7	7	7
			4	7	7
				5	7
					6

We'll leave it to you to figure out how to construct this PIM in **MARK** – all you need to do is modify the logic used for the last PIM a bit. Again, check your parameter indexing. Since the largest value in the survival PIM is 7, the first value in the recapture PIM must be at least 8.

Go ahead and run the model, name it ' $\Phi(t/.), p(t)$ ', and add the results to the browser. There should now be 4 model results in the browser:



Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
$\{\phi(t/.), p(t)\}$	5053.8746	0.0000	0.67789	1.0000	13	133.0730
$\{\phi(t/t), p(t)\}$	5055.3628	1.4882	0.32211	0.4752	16	128.5015
$\{\phi(./.), p(t)\}$	5175.4319	121.5573	0.00000	0.0000	8	264.7031
$\{\phi(t), p(t)\}$	5301.6450	247.7704	0.00000	0.0000	11	384.8765

We see from the results browser that the CJS model is inappropriate for these data – any of the models with age-structure fit better, by any criterion you want to use (no pun intended). In fact, you might want to run these data through **RELEASE** – by looking carefully at the results from TEST 3, you should see some clue as to the 'sources of lack of fit to the CJS model' – in this case, the source being significant age structure. (By the way, you might be – or should be – asking yourselves about doing a GOF test for the general model, to derive an estimate of \hat{c} . The general model in this case (being the model with the most parameters) is model $\{\phi_{t/t} p_t\}$. While we could use the bootstrap, median- \hat{c} or Fletcher- \hat{c} here, there is no need, since the data are simulated under this model with $\hat{c} = 1.0$).

Among the age models, the model with time-dependence in the first age class, but constant survival among the 'adults' has the lowest AIC value, and is a little over twice as well supported by the data than the next best model (the model with time-dependence in both age classes). The LRT comparison of these 2 best models shows that the difference in fit is not significant ($\chi^2_3 = 4.57, P > 0.2$). In other words, the addition of time-dependence to the 'adult' class did not significantly improve the fit – thus we select the more parsimonious model with constant survival in this age class.

Now, before we go much further, how did **MARK** count the parameters for these models? Again, although **MARK** does very well in parameter counting, it is always a good idea to know yourself how many potential parameters there are – if for no other reason than to be able to spot discrepancies between the 'potential' and 'actual' number of parameters estimated given the data. Let's consider the model with the most parameters – model $\{\phi_{t/t} p_t\}$, with time-dependence in both age-classes.

How many individually identifiable parameters are in this model? One *ad hoc* approach to counting parameters, based on 'saturated histories' (histories for individuals encountered at each occasion), was introduced in the Addendum to chapter 4. The saturated capture histories for model $\{\phi_{t/t} p_t\}$, and their associated probability functions, are shown at the top of the next page. Make sure you see how these functions were derived from the survival and recapture matrices.

capture history	probability
1111111	$\varphi_1 p_2 \varphi_7 p_3 \varphi_8 p_4 \varphi_9 p_5 \varphi_{10} p_6 \varphi_{11} p_7$
0111111	$\varphi_2 p_3 \varphi_8 p_4 \varphi_9 p_5 \varphi_{10} p_6 \varphi_{11} p_7$
0011111	$\varphi_3 p_4 \varphi_9 p_5 \varphi_{10} p_6 \varphi_{11} p_7$
0001111	$\varphi_4 p_5 \varphi_{10} p_6 \varphi_{11} p_7$
0000111	$\varphi_5 p_6 \varphi_{11} p_7$
0000011	$\varphi_6 p_7$

Now, of course, the next step is to determine which of the parameters in this table are individually identifiable. By now you may have realized that the 'critical' part of this process typically involves looking at the terminal products. In this case, we have 2 different products: $\varphi_{11} p_7$ and $\varphi_6 p_7$. Are these β terms? If you think about it for a moment, you might realize that the answer is 'yes' – since p_7 is not identifiable. Thus, 16 identifiable parameters. This model corresponds to Table 7D in Lebreton *et al.* (1992). They note that for this model, the number of identifiable parameters is given as $(3k - 5)$, where k is the number of occasions. Since $k = 7$ in this example, we see that $(3k - 5) = (21 - 5) = 16$ parameters. Which is exactly what **MARK** gives us in the results browser.

What about the 2 other age models – $\{\varphi_{t/}, p_t\}$ and $\{\varphi_{t/}, p_t\}$? Start with the first one – since both adult age classes have constant survival, the saturated capture-histories and their corresponding probability statements are:

capture history	probability
1111111	$\varphi_1 p_2 \varphi_2 p_3 \varphi_2 p_4 \varphi_2 p_5 \varphi_2 p_6 \varphi_2 p_7$
0111111	$\varphi_1 p_3 \varphi_2 p_4 \varphi_2 p_5 \varphi_2 p_6 \varphi_2 p_7$
0011111	$\varphi_1 p_4 \varphi_2 p_5 \varphi_2 p_6 \varphi_2 p_7$
0001111	$\varphi_1 p_5 \varphi_2 p_6 \varphi_2 p_7$
0000111	$\varphi_1 p_6 \varphi_2 p_7$
0000011	$\varphi_1 p_7$

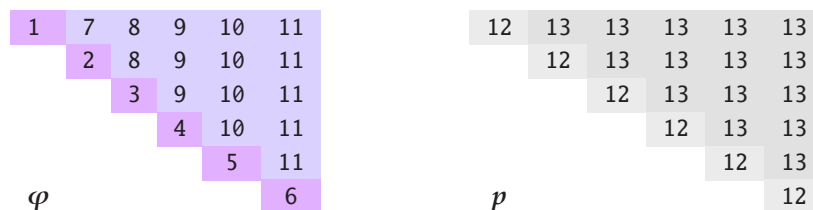
Make sure you can derive the probability statements for this model yourself. How many estimable parameters? **MARK** tells us there are 8 estimable parameters for this model. Is this consistent with the probability statements? There are clearly 2 potential survival parameters (φ_1 and φ_2), and 6 potential recapture parameters (p_2 to p_7) – 8 total potential parameters. Are they all estimable? If you recall from earlier chapters where we deal with models where one or more parameters were held constant, in such cases, there are generally no confounded terminal p and φ since some parameters are estimable because other parameters are held constant across occasions. This is the case here – **all** 8 parameters are estimable, so **MARK** is correct.

What about the other model – model $\{\varphi_{t/}, p_t\}$ – which has both a constant survival (the 'adult' age-class) and time-dependent survival term (the 'juvenile' age class)? Are all terms estimable? Yes! How many? 13 – exactly what **MARK** tells us: 7 survival parameters, and 6 recapture parameters. Why are φ_6 and p_7 separately estimable? Simply because one of the elements (p_7) is estimable from earlier cohorts.

So, we see that **MARK** has correctly calculated the number of estimable parameters for all 3 age models. However, the only way to know if **MARK** is 'correct' is to know how to count the parameters yourself. It can be a somewhat laborious process (admittedly), but it is probably worth the effort.

7.2. Constraining an age model: marked as young only

Constraining a simple age model where the encounter data from individuals marked as young only is not much more difficult than constraining a model without age structure, but there are a few things you need to keep in mind. We'll start by looking at how we would construct the design matrix to correspond to the PIM for survival for the situation with 7 occasions, with 2 age classes and time-dependence in each age class. We'll analyze some simulated data, which are contained in **age_2class.inp**, with the following parameter structure for φ and p , respectively:



So, 2 age-classes for survival, with full time-dependence in each age class, while for the encounter probability, 2 age-classes, but no time variation within each age class.

Here are the results of fitting this model (based on PIMs), using a logit link, to the data:

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
$\{\phi(\tau)\text{p}(J) - \text{PIM} - \text{logit}\}$	8058.3375	0.0000	0.32598	1.0000	13	101.666

Now, the linear models/design matrix approach - we'll focus mainly on modeling survival. Based on the number of indexed parameters in the survival PIM (above), we know that our design matrix for survival will need to have 11 rows and 11 columns.

What does the linear model look like? Again, writing out the linear model is often the easiest place to start. In this case we see that over a given time interval, we have, in effect, 2 kinds of individuals: *juveniles* (individuals in their first year after marking), and *adults* (individuals at least 2 years after marking).

Thus, for a given TIME interval, there are 2 groups: juvenile and adult. If we call this group effect AGE, then we can write out our linear model as

$$\begin{aligned} \text{logit}(\hat{\varphi}) &= \text{AGE} + \text{TIME} + \text{AGE.TIME} \\ &= \beta_1 + \beta_2(\text{AGE}) + \beta_3(\mathbf{t}_1) + \beta_4(\mathbf{t}_2) + \beta_5(\mathbf{t}_3) + \beta_6(\mathbf{t}_4) + \beta_7(\mathbf{t}_5) \\ &\quad + \beta_8(\text{AGE} \cdot \mathbf{t}_7) + \beta_9(\text{AGE} \cdot \mathbf{t}_3) + \beta_{10}(\text{AGE} \cdot \mathbf{t}_4) + \beta_{11}(\text{AGE} \cdot \mathbf{t}_5). \end{aligned}$$

Whoa – wait a minute! How did you get this linear model from the survival PIM? If you look closely at the equation, you'll see that there is no (AGE*t₁) interaction term. What's going on here?

OK. . .step by step. The first term (β_1) is the intercept. Easy enough. The second term (β_2) corresponds to the AGE effect (where in this case we have 2 age classes, juvenile and adult – think of the different ages as different levels of a grouping factor AGE). Thus, one ‘slope’ (beta term), since the 2 levels of the AGE effect require only 1 column of dummy variables.

The next 5 terms ($\beta_3 \rightarrow \beta_7$) correspond to the first 5 levels of TIME. Recall that there are 6 intervals, but that we need only 5 columns to code for these intervals (since if $\tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5 = 0$, then the time interval must be 6). Thus, only 5 ‘slopes’ (beta terms) coding for the 6 different time intervals.

Now, the potentially ‘tricky’ part – the interaction terms. If you look carefully – very carefully – at the linear model equation

$$\begin{aligned}\text{logit}(\hat{\phi}) &= \text{AGE} + \text{TIME} + \text{AGE}.\text{TIME} \\ &= \beta_1 + \beta_2(\text{AGE}) + \beta_3(\text{t}_1) + \beta_4(\text{t}_2) + \beta_5(\text{t}_3) + \beta_6(\text{t}_4) + \beta_7(\text{t}_5) \\ &\quad + \beta_8(\text{AGE}.\text{t}_2) + \beta_9(\text{AGE}.\text{t}_3) + \beta_{10}(\text{AGE}.\text{t}_4) + \beta_{11}(\text{AGE}.\text{t}_5).\end{aligned}$$

you’ll notice that it did **not** include a term for an (AGE.T₁) interaction. Why? You need to think a bit carefully here. Remember, we’re treating the two age classes as (in effect) different groups. If you look back at the examples in Chapter 4, you’ll notice that, in those cases, the groups being compared (male vs. female, good vs. poor) were ‘temporally symmetrical’ – i.e., had the same number of sampling occasions for both groups. In our present age-structured example, however, this is **not** the case. Look closely at the survival PIM –

1	7	7	7	7	7
2	7	7	7	7	
3	7	7	7		
4	7	7			
5	7				
6					

Note that we have 6 parameters for the juvenile age class (1 → 6, corresponding to the first 6 intervals of the study) but only 5 parameters for the adult age class (7 → 11, corresponding to intervals 2 → 6).

If you look closely, you’ll see that we don’t have an ‘adult’ estimate over the first interval for the first cohort, since there were no ‘known’ (marked) adults at that point in the study. Thus, in the first time interval, there are no adults, and thus, there is no logical interaction of the AGE and TIME factors in this interval (since there is only one age class!). In other words, no (AGE.t₁) term. The first time interval for which there is a potential interaction of AGE and TIME is interval 2 (look at the PIM again to confirm this for yourself). If you don’t see the connection between the PIM, and the linear model, take some time now to work through everything, to make sure that you do. It’s a very important concept.

Here is the design matrix for our age model, with time-dependence in both age classes for survival (we’ll use a simple identity structure for p):

B1 int	B2 age	B3 t1	B4 t2	B5 t3	B6 t4	B7 t5	B8 at2	B9 at3	B10 at4	B11 at5	Pam	B12 p2	B13 p3
1	1	1	0	0	0	0	0	0	0	0	1:Phi	0	0
1	1	0	1	0	0	0	1	0	0	0	2:Phi	0	0
1	1	0	0	1	0	0	0	1	0	0	3:Phi	0	0
1	1	0	0	0	1	0	0	0	1	0	4:Phi	0	0
1	1	0	0	0	0	1	0	0	0	1	5:Phi	0	0
1	1	0	0	0	0	0	0	0	0	0	6:Phi	0	0
1	0	0	1	0	0	0	0	0	0	0	7:Phi	0	0
1	0	0	0	1	0	0	0	0	0	0	8:Phi	0	0
1	0	0	0	0	1	0	0	0	0	0	9:Phi	0	0
1	0	0	0	0	0	1	0	0	0	0	10:Phi	0	0
1	0	0	0	0	0	0	0	0	0	0	11:Phi	0	0
0	0	0	0	0	0	0	0	0	0	0	12:p	1	0
0	0	0	0	0	0	0	0	0	0	0	13:p	0	1

Make sure you see the connection between the linear model, and this design matrix. Note in particular

that (i) there are 6 rows corresponding to intervals $1 \rightarrow 6$ for juveniles, with 5 rows corresponding to intervals $2 \rightarrow 6$ for adults, and (ii) that the time indexing (along the diagonal) starts at interval 2 for the adults, and for the interaction columns.

Go ahead and run this model, and add the results to the browser (below). As expected, the fits of both the PIM- and DM-based models to the data are identical:

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
{phi(t)t)p(.). - PIM - logit}	8058.3375	0.0000	0.32598	1.0000	13	101.6660
{phi(t)t)p(.). - DM}	8058.3375	0.0000	0.32598	1.0000	13	101.6660

It is important to understand that here, we are using a ‘common intercept’ for both age classes, since we might want to evaluate the fit of a model where (say) survival of adults is additive to the survival of the juveniles (i.e., adult survival is model as a linear function of juvenile survival plus an additive constant).

This we could implement in our design matrix simply by deleting the ‘interaction terms’ (columns B8 \rightarrow B11), leaving:

B1: phi-int	B2: age	B3: t1	B4: t2	B5: t3	B6: t4	B7: t5	Parm	B8: p2	B9: p3
1	1	1	0	0	0	0	1:Phi	0	0
1	1	0	1	0	0	0	2:Phi	0	0
1	1	0	0	1	0	0	3:Phi	0	0
1	1	0	0	0	1	0	4:Phi	0	0
1	1	0	0	0	0	1	5:Phi	0	0
1	1	0	0	0	0	0	6:Phi	0	0
1	0	0	1	0	0	0	7:Phi	0	0
1	0	0	0	1	0	0	8:Phi	0	0
1	0	0	0	0	1	0	9:Phi	0	0
1	0	0	0	0	0	1	10:Phi	0	0
1	0	0	0	0	0	0	11:Phi	0	0
0	0	0	0	0	0	0	12:p	1	0
0	0	0	0	0	0	0	13:p	0	1

which corresponds to the additive model for ‘AGE+TIME’:

$$\begin{aligned}\text{logit}(\hat{\phi}) &= \text{AGE} + \text{TIME} \\ &= \beta_1 + \beta_2(\text{AGE}) + \beta_3(t_1) + \beta_4(t_2) + \beta_5(t_3) + \beta_6(t_4) + \beta_7(t_5).\end{aligned}$$

Go ahead and run this model, and add the results to the browser:

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
{phi(t)t)p(.). - PIM - logit}	8058.3375	0.0000	0.48363	1.0000	13	101.6660
{phi(t)t)p(.). - DM}	8058.3375	0.0000	0.48363	1.0000	13	101.6660
{phi(t+t)p(.). - DM}	8063.7327	5.3952	0.03258	0.0674	9	115.1011

But, suppose for some reason we were not interested in an ‘additive’ model. Suppose we simply wanted to treat ‘juveniles’ and ‘adults’ as ‘completely separate’ groups, each allowed to vary uniquely, independent of the other. In fact, this is exactly what is represented by the full interaction model we started with (above).

But, if we have no interest in an ‘additive’ model, we could have built such a model without using a ‘common intercept’ approach – treating ‘juvenile’ and ‘adult’ survival as (in effect) separate parameters, rather than treating them as separate ‘groups’ for a common parameter.* We do this by coding each ‘age class’ with its own intercept, as follows:

B1: int-juv	B2: t1	B3: t2	B4: t3	B5: t4	B6: t5	Parm	B7: int-adt	B8: t2	B9: t3	B10: t4	B11: t5	B12: p2	B13: p3
1	1	0	0	0	0	1:Phi	0	0	0	0	0	0	0
1	0	1	0	0	0	2:Phi	0	0	0	0	0	0	0
1	0	0	1	0	0	3:Phi	0	0	0	0	0	0	0
1	0	0	0	1	0	4:Phi	0	0	0	0	0	0	0
1	0	0	0	0	1	5:Phi	0	0	0	0	0	0	0
1	0	0	0	0	0	6:Phi	0	0	0	0	0	0	0
0	0	0	0	0	0	7:Phi	1	0	0	0	0	0	0
0	0	0	0	0	0	8:Phi	1	0	1	0	0	0	0
0	0	0	0	0	0	9:Phi	1	0	0	1	0	0	0
0	0	0	0	0	0	10:Phi	1	0	0	0	1	0	0
0	0	0	0	0	0	11:Phi	1	0	0	0	0	0	0
0	0	0	0	0	0	12:p	0	0	0	0	0	1	0
0	0	0	0	0	0	13:p	0	0	0	0	0	0	1

Several things to note, here. First, the total number of columns in the DM does not change – we still have 11 columns coding for the survival parameters, plus 2 for the encounter parameters. Second, note that ‘juvenile’ and ‘adult’ survival (parameters 1 → 6 and 7 → 11, respectively), each have their own intercept. Thus, each set of parameters is modeled as if (in effect) it was its own parameter, with its own linear model:

$$\begin{aligned}\text{logit}(\hat{\phi}_{\text{juv}}) &= \text{TIME} \\ &= \beta_1 + \beta_2(t_1) + \beta_3(t_2) + \beta_4(t_3) + \beta_5(t_4) + \beta_6(t_5), \\ \text{logit}(\hat{\phi}_{\text{adt}}) &= \text{TIME} \\ &= \beta_7 + \beta_8(t_2) + \beta_9(t_3) + \beta_{10}(t_4) + \beta_{11}(t_5) + \beta_{12}(t_6).\end{aligned}$$

Go ahead and run this model, and add the results to the browser:

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
{phi(t)p(.)} - PIM - logit	8058.3375	0.0000	0.32598	1.0000	13	101.6660
{phi(t)p(.)} - DM	8058.3375	0.0000	0.32598	1.0000	13	101.6660
{phi(t)p(.)} - DM - separate intercept approach	8058.3375	0.0000	0.32598	1.0000	13	101.6660
{phi(t)p(.)} - DM	8063.7327	5.3952	0.02196	0.0674	9	115.1011

As expected, the model deviance and likelihood for this ‘separate intercept’ models is identical to the full interaction model built using with the PIM or DM approach. Beyond increasing your overall understanding of the relationship between the linear models and the design matrix, this ‘separate intercepts’ approach can be useful in complex models where setting up the main effect interactions can be ‘challenging’ (as we will see in the next section).

* This idea of thinking of the two age classes as ‘separate parameters’ will show up again – at least structurally – when we consider abundance estimation from closed populations – Chapter 15.

To further test your understanding of simple age models and the design matrix, let's consider a couple of more examples. First, consider a model with 2 age-classes for apparent survival, with time-dependence for the first (juvenile) age-class only (adult survival is held constant over time) (model $\varphi_{t/}$). This is a very common structure to consider, especially for long(er)-lived organisms where juvenile survival often varies strongly over time, with comparatively time-invariant survival for adults.

The PIM structure for φ for this model is shown below:

1	7	7	7	7	7
2	7	7	7	7	
3	7	7	7		
4	7	7			
5	7				
6					

Let's first build this model $\{\varphi_{t/}, p_{./}\}$ using PIMs, and add the results to the browser:

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
$\{\varphi_{t/}p_{./} - \text{PIM} - \text{logit}\}$	8058.3375	0.0000	0.49996	1.0000	13	101.6660
$\{\varphi_{t/}p_{./} - \text{DM}\}$	8058.3375	0.0000	0.49996	1.0000	13	101.6660
$\{\varphi_{t/}p_{./} - \text{PIM}\}$	8075.7988	17.4613	0.00008	0.0002	9	127.1672

Now, let's try building this same model, but using a DM approach. Start by retrieving model $\{\varphi_{t/}, p_{./}\}$ from the browser – the DM for that model is shown below:

B1 int	B2 age	B3 t1	B4 t2	B5 t3	B6 t4	B7 t5	B8 at2	B9 at3	B10 at4	B11 at5	Parm	B12 p2	B13 p3
1	1	1	0	0	0	0	0	0	0	0	1:Phi	0	0
1	1	0	1	0	0	0	1	0	0	0	2:Phi	0	0
1	1	0	0	1	0	0	0	1	0	0	3:Phi	0	0
1	1	0	0	0	1	0	0	0	1	0	4:Phi	0	0
1	1	0	0	0	0	1	0	0	0	1	5:Phi	0	0
1	1	0	0	0	0	0	0	0	0	0	6:Phi	0	0
1	0	0	1	0	0	0	0	0	0	0	7:Phi	0	0
1	0	0	0	1	0	0	0	0	0	0	8:Phi	0	0
1	0	0	0	0	1	0	0	0	0	0	9:Phi	0	0
1	0	0	0	0	0	1	0	0	0	0	10:Phi	0	0
1	0	0	0	0	0	0	0	0	0	0	11:Phi	0	0
0	0	0	0	0	0	0	0	0	0	0	12:p	1	0
0	0	0	0	0	0	0	0	0	0	0	13:p	0	1

In this design matrix we have time-variation for both juvenile and adult age classes.

However, here we are trying to build model $\{\varphi_{t/}, p_{./}\}$ – time-variation in juvenile survival only. So, we need to modify the DM such that survival is held constant for the adult age class. It might help to consider the linear model for the TIME factor, for the two age classes separately. For juveniles, with time variation, the linear model would consist of 5 β terms, coding for the 6 time intervals. In contrast, for adults, survival is constant over time. Recall from chapter 6 that such a constant ('dot') model can be modeled by using a simple 'intercept' – i.e., a single column of 1's.

Keeping this in mind, here is the DM modified to represent model $\{\varphi_{t/}, p_{./}\}$:

B1: int	B2: age	B3: t1	B4: t2	B5: t3	B6: t4	B7: t5	Parm
1	1	1	0	0	0	0	1:Phi
1	1	0	1	0	0	0	2:Phi
1	1	0	0	1	0	0	3:Phi
1	1	0	0	0	1	0	4:Phi
1	1	0	0	0	0	1	5:Phi
1	1	0	0	0	0	0	6:Phi
1	0	0	0	0	0	0	7:Phi
1	0	0	0	0	0	0	8:Phi
1	0	0	0	0	0	0	9:Phi
1	0	0	0	0	0	0	10:Phi
1	0	0	0	0	0	0	11:Phi

Note that we no longer have any interaction of age with time (since there is no longer time variation for the adults), and that we have eliminated all the time indexing for the adults (parameters $7 \rightarrow 11$).

If we run this model, and add the results to the browser (shown below), we see that it yields results which are identical to that from the model built using PIMs (which indicates that the two models are equivalent; i.e., that our DM is correct).

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
$\{\phi(t)p_{./} - \text{PIM} - \text{logit}\}$	8058.3375	0.0000	0.49992	1.0000	13	101.6660
$\{\phi(t)p_{./} - \text{DM}\}$	8058.3375	0.0000	0.49992	1.0000	13	101.6660
$\{\phi(t)p_{./} - \text{PIM}\}$	8075.7988	17.4613	0.00008	0.0002	9	127.1672
$\{\phi(t)p_{./} - \text{DM}\}$	8075.7988	17.4613	0.00008	0.0002	9	127.1672

As per the previous example, we could also have used the ‘separate intercept’ approach for this model, using the following:

B1: int-juv	B2: t1	B3: t2	B4: t3	B5: t4	B6: t5	Parm	B7: int-adt	B8: p2	B9: p3
1	1	0	0	0	0	1:Phi	0	0	0
1	0	1	0	0	0	2:Phi	0	0	0
1	0	0	1	0	0	3:Phi	0	0	0
1	0	0	0	1	0	4:Phi	0	0	0
1	0	0	0	0	1	5:Phi	0	0	0
1	0	0	0	0	0	6:Phi	0	0	0
0	0	0	0	0	0	7:Phi	1	0	0
0	0	0	0	0	0	8:Phi	1	0	0
0	0	0	0	0	0	9:Phi	1	0	0
0	0	0	0	0	0	10:Phi	1	0	0
0	0	0	0	0	0	11:Phi	1	0	0
0	0	0	0	0	0	12:p	0	1	0
0	0	0	0	0	0	13:p	0	0	1

In fact, for this particular model, this approach might actually be preferred, since it conveniently separates (literally) juvenile survival, which you might model as a function of some temporally varying *environmental* covariate, from adult survival, which is held constant over time (although you could model

both juvenile and adult survival as a function of an *individual* covariate, even using this structure – individual covariates are discussed in Chapter 11).

As shown below, the two approaches to building model $\{\varphi_{t/}, p_{./}\}$ yield identical model fits to the data:

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
$\{\phi(t)p(.)/ - \text{PIM} - \text{logit}\}$	8058.3375	0.0000	0.49988	1.0000	13	101.6660
$\{\phi(t)p(.)/ - \text{DM}\}$	8058.3375	0.0000	0.49988	1.0000	13	101.6660
$\{\phi(t)p(.)/ - \text{PIM}\}$	8075.7988	17.4613	0.00008	0.0002	9	127.1672
$\{\phi(t)p(.)/ - \text{DM}\}$	8075.7988	17.4613	0.00008	0.0002	9	127.1672
$\{\phi(t)p(.)/ - \text{DM} - \text{separate intercept approach}\}$	8075.7988	17.4613	0.00008	0.0002	9	127.1672

A final example concerning 2-age models for individuals marked as young only. Consider now a model with 2 age-classes, with time-dependence in each age class. Here, however, we have some reason to believe that there is a linear change in survival over time in both age classes (i.e., that survival is varying systematically – increasing or decreasing – over time). We want perhaps to compare the slopes of the linear trend in survival between young and old individuals. If the slopes are not significantly different, we could test against a model with a common slope for both age classes.

From the preceding example, we know that for ‘juveniles’, the parameters of interest are $1 \rightarrow 6$, and for ‘adults’, the parameters are $7 \rightarrow 11$. But, how do we handle the vector of increasing values to handle the trend? Recall from Chapter 6 that we could ‘model’ trend simply by specifying an ordinal sequence of numbers in the design matrix. But, should we use $1 \rightarrow 6$ for juveniles, and $2 \rightarrow 6$ for adults (which corresponds to where the overlap occurs), or is it equivalent to use $1 \rightarrow 6$ and $1 \rightarrow 5$ respectively?

As it turns out, it doesn’t matter at all to **MARK** – at least not mechanically (i.e., **MARK** will run just fine in either case). However, note that in the second case you would, in effect, be coding each time step differently for each group, which makes the intercepts no longer comparable. Thus, if you choose to accommodate the overlap, the first coding scheme ($1 \rightarrow 6$ for juveniles, and $2 \rightarrow 6$ for adults in this example) is preferred. Another possibility, of course, is to drop the first juvenile parameter altogether from our constraint. In other words, to include only those parameters which ‘overlap’ (i.e., $2 \rightarrow 6$ and $7 \rightarrow 11$).

For our trend analysis, we’ll use the former approach, which includes survival for the first age class over the first interval. In this case, the design matrix for the survival elements would look like:

	B1	B2	B3	B4	Parm
1	1	1	1	1	1:Phi
1	1	2	2	2	2:Phi
1	1	3	3	3	3:Phi
1	1	4	4	4	4:Phi
1	1	5	5	5	5:Phi
1	1	6	6	6	6:Phi
1	0	2	2	0	7:Phi
1	0	3	3	0	8:Phi
1	0	4	4	0	9:Phi
1	0	5	5	0	10:Phi
1	0	6	6	0	11:Phi

All you need to do now is run **MARK**, and apply the constraint to the underlying model (2 age-classes with time-dependence for survival, simple time-dependence for recaptures). And, as discussed in detail in Chapter 6, by varying which columns of the design matrix you use in the constraint, you

can test hypotheses concerning equivalence of slope between age classes, equivalence of intercepts, and so forth. You could also test for additivity (parallelism) in survival for a time-dependent model. Everything covered in Chapter 6 in terms of constraining an underlying CJS model applies equally well to age models (or cohort models, or **any** models).

An important ‘nuance’ – continuous covariates

In the preceding, we considered the construction of the interaction terms for 2 types of models: one where ‘age’ and ‘time’ were both (classification) factors, and one where ‘age’ interacted with ‘time’ constrained to follow a linear trend over time. In the first case, we considered the complication that there is no logical interaction of ‘age’ and ‘time’ for the first interval, since only the first juvenile age class is present in our marked sample over the first interval (i.e., no adults). In this case, we only included time intervals 2 → 6 in the interaction of ‘age’ and ‘time’, as shown (again) in the following DM:

B1 int	B2 age	B3 t1	B4 t2	B5 t3	B6 t4	B7 t5	B8 a12	B9 a13	B10 a14	B11 a15	Pam	B12 p2	B13 p3
1	1	1	0	0	0	0	0	0	0	0	1:Phi	0	0
1	1	0	1	0	0	0	1	0	0	0	2:Phi	0	0
1	1	0	0	1	0	0	0	1	0	0	3:Phi	0	0
1	1	0	0	0	1	0	0	0	1	0	4:Phi	0	0
1	1	0	0	0	0	1	0	0	0	1	5:Phi	0	0
1	1	0	0	0	0	0	0	0	0	0	6:Phi	0	0
1	0	0	1	0	0	0	0	0	0	0	7:Phi	0	0
1	0	0	0	1	0	0	0	0	0	0	8:Phi	0	0
1	0	0	0	0	1	0	0	0	0	0	9:Phi	0	0
1	0	0	0	0	0	1	0	0	0	0	10:Phi	0	0
1	0	0	0	0	0	0	0	0	0	0	11:Phi	0	0
0	0	0	0	0	0	0	0	0	0	0	12p	1	0
0	0	0	0	0	0	0	0	0	0	0	13p	0	1

In the second case, where ‘time’ was constrained to follow a linear trend within each age class, we used the following DM:

B1	B2	B3	B4	Pam
1	1	1	1	1:Phi
1	1	2	2	2:Phi
1	1	3	3	3:Phi
1	1	4	4	4:Phi
1	1	5	5	5:Phi
1	1	6	6	6:Phi
1	0	2	0	7:Phi
1	0	3	0	8:Phi
1	0	4	0	9:Phi
1	0	5	0	10:Phi
1	0	6	0	11:Phi

While this seems reasonable enough at first glance, you might now wonder why we included the interaction for the first time interval, since there are only juveniles present during that interval. For example, why didn’t we use something like the DM shown at the top of the next page? In other words, why didn’t we put a ‘0’ in the first row of the interaction column?

B1:	B2:	B3:	B4:	Parm	
1	1	1	0	1:Phi	0
1	1	2	2	2:Phi	0
1	1	3	3	3:Phi	0
1	1	4	4	4:Phi	0
1	1	5	5	5:Phi	0
1	1	6	6	6:Phi	0
1	0	2	0	7:Phi	0
1	0	3	0	8:Phi	0
1	0	4	0	9:Phi	0
1	0	5	0	10:Phi	0
1	0	6	0	11:Phi	0

As it turns out, if we try this approach, **MARK** will give us the wrong answer. How do we know this?

Let's first start by considering a different approach to modeling trend within each age classes, with interaction between the age classes. Rather than use a common intercept, we'll use a DM where each age class has it's own intercept (as per several of the examples presented earlier in this section):

B1:	B2:	B3:	B4:	Parm	
1	1	0	0	1:Phi	0
1	2	0	0	2:Phi	0
1	3	0	0	3:Phi	0
1	4	0	0	4:Phi	0
1	5	0	0	5:Phi	0
1	6	0	0	6:Phi	0
0	0	1	2	7:Phi	0
0	0	1	3	8:Phi	0
0	0	1	4	9:Phi	0
0	0	1	5	10:Phi	0
0	0	1	6	11:Phi	0

If we run this model, we see that in fact it yields the same deviance and parameter estimates as the model we fit originally, using a common intercept (i.e., the DM at the bottom of the preceding page).

What is important here is that in this case, 'time' is no longer a simple, unconstrained factor, but is being constrained as a linear function – in this case, constrained to follow a linear trend. When 'time' is an unconstrained factor, we need to only include interactions where both 'time' and 'age' occur. However, when 'time' is constrained to be a linear function of something (say, a trend, or some environmental covariate), then we need to include all time steps in the interaction.

We can further demonstrate this by means of another example. Suppose that instead of constraining variation in 'time' to follow a trend, we constrain variation to be a linear function of some environmental covariate (say, average temperature). Suppose that the average temperature over each of 6 intervals is: {10, 14, 22, 11, 15, 17}.

The DM using a separate intercept for each age class, where survival within a given age class is constrained to be a function of the average temperature, shown at the top of the next page. The model deviance for this DM fit to the data (with full time-dependence in p) is 5,102.437.

B1:	B2:	B3:	B4:	Parm	
1	10	0	0	1:Phi	0
1	14	0	0	2:Phi	0
1	22	0	0	3:Phi	0
1	11	0	0	4:Phi	0
1	15	0	0	5:Phi	0
1	17	0	0	6:Phi	0
0	0	1	14	7:Phi	0
0	0	1	22	8:Phi	0
0	0	1	11	9:Phi	0
0	0	1	15	10:Phi	0
0	0	1	17	11:Phi	0

Now, if we try a DM with separate intercepts for each age class, and leave out the interaction for the first time interval for juveniles

B1:	B2:	B3:	B4:	Parm
1	1	10	0	1:Phi
1	1	14	14	2:Phi
1	1	22	22	3:Phi
1	1	11	11	4:Phi
1	1	15	15	5:Phi
1	1	17	17	6:Phi
1	0	14	0	7:Phi
1	0	22	0	8:Phi
1	0	11	0	9:Phi
1	0	15	0	10:Phi
1	0	17	0	11:Phi

yields a model deviance of 5,088.601, which is not the same. And, clearly, then, it isn't the same model we constructed initially using separate intercepts for each age class.

However, if we include the first interval for juveniles in the interaction

B1:	B2:	B3:	B4:	Parm
1	1	10	10	1:Phi
1	1	14	14	2:Phi
1	1	22	22	3:Phi
1	1	11	11	4:Phi
1	1	15	15	5:Phi
1	1	17	17	6:Phi
1	0	14	0	7:Phi
1	0	22	0	8:Phi
1	0	11	0	9:Phi
1	0	15	0	10:Phi
1	0	17	0	11:Phi

the model deviance is 5,102.437, identical to the model we fitting using a DM with separate intercepts.

The reason for the difference here – needing to include the covariate interaction for all time intervals – is because we need to fully specify the linear relationship between the parameter and the covariate. Leaving out one covariate for a particular age class, and replacing it with a 0, would change the nature of the linear model for the covariate.

In contrast, for 'time' as a factor, there is no underlying linear model for time. This is a subtle point of distinction, and one you need to think about a bit, whenever you have a 'grouping' factor (like 'age') with unequal number of time intervals (say, between juveniles and adults).

7.2.1. DM with > 2 age classes: 'ugly' interaction terms

As noted in the preceding examples involving 2 age classes, the key consideration to building the design matrix for age models in general are (i) keeping track of the asymmetry of which age classes in the marked sample are represented at each occasion, and (ii) how to handle the interactions among age classes. These challenges are compounded for analysis involving > 2 age classes. Here, again, we consider only the relatively simple situation where the encounter data are from individuals initially marked as young.

Consider the following PIM, for (say) apparent survival, ϕ :

1	7	12	13	14	15
	2	8	13	14	15
		3	9	14	15
			4	10	15
				5	11
					6

Here, we have specified 3 age classes with time-dependence within each age class: parameters 1 \rightarrow 6 for the first age class (for the first 6 time intervals), parameters 7 \rightarrow 11 for the second age class (for time intervals 2 \rightarrow 6), and parameters 12 \rightarrow 15 for the third and final age class (for time intervals 3 \rightarrow 6). We might call this model $\{\phi_{t/t/t}\}$.

With > 2 age classes, we'll need to have > 1 β parameter (slope) in the linear model to code for 'age group': $(k - 1) = (3 - 1) = 2$ parameters, in fact. In addition, for 7 occasions (6 time intervals) we need $(k - 1) = (6 - 1) = 5$ parameters for time.

Thus, our linear model starts with the following structure, with AGE_i represent the coding for AGE group, and t_i representing the coding for TIME:

$$\text{logit}(\hat{\phi}) = \beta_1 + \beta_2(AGE_1) + \beta_3(AGE_2) + \beta_4(t_1) + \beta_5(t_2) + \beta_6(t_3) + \beta_7(t_4) + \beta_8(t_5) \\ + \text{'numerous interaction terms'}.$$

While coding the main effects (AGE and TIME) is straightforward at this point, the remaining challenge, then, concerns the interaction terms, and we'll need to be somewhat careful in constructing them. Since there is 1 intercept (β_1), followed by 2 columns of dummy variables for 'AGE' (β_2 and β_3), and then 5 columns for 'TIME' ($\beta_4 \rightarrow \beta_8$), then you might initially think we'd need $(2 \times 5) = 10$ columns for the interactions, and thus $(1 + 7 + 10) = 18$ total parameters (i.e., β terms) in our linear model.

However, if you lookback at the PIM (above), you might realize that this conclusion would be incorrect. How many parameters are there in the PIM? Clearly, there are 15 parameters. As such, the linear model for this parameter structure can only have 15 parameters, not 18. In fact, the linear model must have exactly 15 terms – one for each parameter in the PIM. This is your 'check' for your linear model – it must have a β parameter from each parameter indexed in the PIM. No more, no less.

So, clearly, our initial conclusion of 10 interaction parameters must be incorrect – but, why? The problem is that when we calculated the number of interaction columns, we neglected to account for

the fact that in an 'age' model, not all interactions between 'age' and 'time' are *plausible*. Have another look at the PIM structure. What are the intervals for which there are interactions between age class 1 individuals (first diagonal) and age class 2 individuals (second diagonal)? In other words, in what intervals might the estimates differ between these 2 age classes?

Consider the following rendition of the underlying survival PIM:

1	7	12	13	14	15
	2	8	13	14	15
		3	9	14	15
			4	10	15
				5	11
					6

Here, the shaded parts of the PIM indicate clearly where these interactions occur – they occur in intervals (2 → 6). There is no interaction of age class 1 and age class 2 individuals in the marked population for interval 1, since there are no age class 2 individuals in interval 1! This is exactly the same situation we saw for the simpler 2 age class model introduced at the start of this section.

Let's add terms to the linear model to reflect these *plausible* interactions between these 2 age classes (i.e., age class 1 and age class 2). We enter a term in the linear model (below, second line of the equation) for each of the intervals for which there is a logical (possible) interaction of these two age classes (intervals 2 → 6); $k = 5$ intervals, 4 β terms):

$$\begin{aligned} \text{logit}(\hat{\varphi}) = & \beta_1 + \beta_2(\text{AGE}_1) + \beta_3(\text{AGE}_2) + \beta_4(\text{T}_1) + \beta_5(\text{T}_2) + \beta_6(\text{T}_3) + \beta_7(\text{T}_4) + \beta_8(\text{T}_5) \\ & + \beta_9(\text{AGE}_1 \cdot \text{T}_2) + \beta_{10}(\text{AGE}_1 \cdot \text{T}_3) + \beta_{11}(\text{AGE}_1 \cdot \text{T}_4) + \beta_{12}(\text{AGE}_1 \cdot \text{T}_5) \\ & + \text{'remaining interaction terms'} \end{aligned}$$

If you looked at the equation closely, you might have noticed that we used the subscript '1' for AGE for these new interaction β terms. Why? Remember that we have 3 age classes here, so we need 2 columns of dummy variables to code for it. If we use

$$\beta_2 = \begin{cases} 1 & \text{if age} = 1 \\ 0 & \text{if other} \end{cases} \quad \beta_3 = \begin{cases} 1 & \text{if age} = 2 \\ 0 & \text{if other} \end{cases}$$

then AGE_{3+} is the reference age class (for individuals age ≥ 2 years), AGE_1 corresponds to age-class 1 year individuals, and AGE_2 corresponds to age-class 2 year individuals. So, for the first interaction, you're multiplying the column for β_2 against the $(5 - 1) = 4$ columns for those time intervals where the interaction between age class 1 and age class 2 occurs, which yields 4 β terms for that interaction.

What about for the interactions between age class 2 and age class 3+ individuals? Again, have a careful look at the re-drawn PIM for survival (below):

1	7	12	13	14	15
	2	8	13	14	15
		3	9	14	15
			4	10	15
				5	11
					6

The shaded parts of the PIM indicate where (i.e., over which time intervals) the interactions of age class 2 and age class 3+ individuals occur – over intervals (3 → 6).

So, as above, we enter a term in the linear model (below, third line of the equation) for each of the intervals for which there is a logical (possible) interaction of these two age classes (in this case, intervals (3 → 6); $k = 4$ intervals, $(k - 1) = (4 - 1) = 3$ separate β terms):

$$\begin{aligned} \text{logit}(\hat{\varphi}) = & \beta_1 + \beta_2(\text{AGE}_1) + \beta_3(\text{AGE}_2) + \beta_4(\text{T}_1) + \beta_5(\text{T}_2) + \beta_6(\text{T}_3) + \beta_7(\text{T}_4) + \beta_8(\text{T}_5) \\ & + \beta_9(\text{AGE}_1 \cdot \text{T}_2) + \beta_{10}(\text{AGE}_1 \cdot \text{T}_3) + \beta_{11}(\text{AGE}_1 \cdot \text{T}_4) + \beta_{12}(\text{AGE}_1 \cdot \text{T}_5) \\ & + \beta_{13}(\text{AGE}_2 \cdot \text{T}_3) + \beta_{14}(\text{AGE}_2 \cdot \text{T}_4) + \beta_{15}(\text{AGE}_2 \cdot \text{T}_5) \end{aligned}$$

For these interactions, you’re multiplying the column for β_3 , reflecting AGE_2 , against the $(4 - 1) = 3$ columns for those time intervals where the interaction between 2 and 3+ occurs, which yields 3 additional β terms for that interaction.

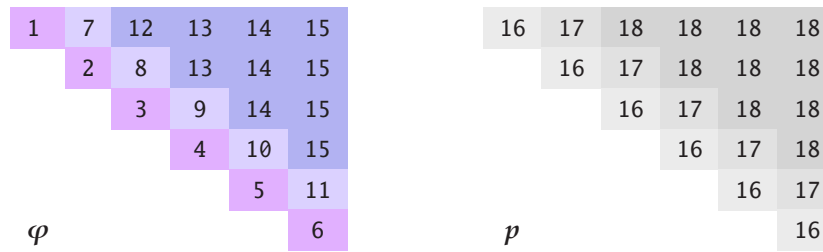
Note that our linear model (above) now has 15 terms ($\beta_1 \rightarrow \beta_{15}$), matching the number of parameters in the underlying PIM. This would imply that in fact, our linear model must be ‘complete’.

It will help to look closely at the dummy variable coding for the different AGE groups in the DM (shown below), which corresponds to the structure of the linear model shown above:

B1 int	B2 age1	B3 age2	B4 t1	B5 t2	B6 t3	B7 t4	B8 t5	B9 a1t2	B10 a1t3	B11 a1t4	B12 a1t5	B13 a2t3	B14 a2t4	B15 a2t5	Pam
1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1:Phi
1	1	0	0	1	0	0	0	1	0	0	0	0	0	0	2:Phi
1	1	0	0	0	1	0	0	0	1	0	0	0	0	0	3:Phi
1	1	0	0	0	0	1	0	0	0	1	0	0	0	0	4:Phi
1	1	0	0	0	0	0	1	0	0	0	1	0	0	0	5:Phi
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	6:Phi
1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	7:Phi
1	0	1	0	0	1	0	0	0	0	0	0	1	0	0	8:Phi
1	0	1	0	0	0	1	0	0	0	0	0	0	1	0	9:Phi
1	0	1	0	0	0	0	1	0	0	0	0	0	0	1	10:Phi
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	11:Phi
1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	12:Phi
1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	13:Phi
1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	14:Phi
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15:Phi

If you look at the DM closely, you’ll recall that a few pages back, we made the arbitrary (in this example) choice that age-class 3+ individuals are coded as ‘0’ for β_2 and ‘0’ for β_3 , and thus – since the product of 0 with anything is still 0 – there are no possible interaction terms possible. So, in fact, the preceding design matrix (above) is the final, finished design matrix, and our linear model is in fact, complete.

As a final check, we’ll demonstrate that we ‘have it right’ empirically. We’ll use some simulated live encounter data (**age_3class.inp**), generated from the ‘true’ parameter structure for φ and p , respectively, shown in the PIMs at the top of the next page.



So, 3 age-classes for survival, with full time-dependence in each age class, while for the encounter probability, 3 age-classes, but no time variation within each age class. We'll call this model $\{\varphi_{t/t/t} p_{./././}\}$.

So, we'll start by building the model using PIMs (based on the above). We'll fit this model using the logit link, and add the results to the browser:

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
$\{\text{phi}(t/t)p(J,J) - \text{logit} - \text{PIM}\}$	5171.2223	0.0000	0.33333	1.0000	18	99.5453

Now, we construct the same model, using the DM we developed on the previous page. For modeling encounter probability, we'll use a simple identity matrix:

B1: int	B2: age1	B3: age2	B4: t1	B5: t2	B6: t3	B7: t4	B8: t5	B9: a1t2	B10: a1t3	B11: a1t4	B12: a1t5	B13: a2t3	B14: a2t4	B15: a2t5	Pam	B16: p2	B17: p3	B18: p4
1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1:Phi	0	0	0
1	1	0	0	1	0	0	0	1	0	0	0	0	0	0	2:Phi	0	0	0
1	1	0	0	0	1	0	0	0	1	0	0	0	0	0	3:Phi	0	0	0
1	1	0	0	0	0	1	0	0	0	1	0	0	0	0	4:Phi	0	0	0
1	1	0	0	0	0	0	1	0	0	0	1	0	0	0	5:Phi	0	0	0
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	6:Phi	0	0	0
1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	7:Phi	0	0	0
1	0	1	0	0	1	0	0	0	0	0	0	1	0	0	8:Phi	0	0	0
1	0	1	0	0	0	1	0	0	0	0	0	0	1	0	9:Phi	0	0	0
1	0	1	0	0	0	0	1	0	0	0	0	0	0	1	10:Phi	0	0	0
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	11:Phi	0	0	0
1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	12:Phi	0	0	0
1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	13:Phi	0	0	0
1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	14:Phi	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15:Phi	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16:p	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17:p	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18:p	0	0	1

As expected, the results are identical between the two models:

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
$\{\text{phi}(t/t)p(J,J) - \text{logit} - \text{PIM}\}$	5171.2223	0.0000	0.33333	1.0000	18	99.5453
$\{\text{phi}(t/t)p(J,J) - \text{DM} - \text{age 3 reference}\}$	5171.2223	0.0000	0.33333	1.0000	18	99.5453

How would the design matrix change if you had made age 1 the reference age class in your dummy variable coding scheme? The change in coding is shown at the top of the next page.

$$\beta_2 = \begin{cases} 1 & \text{if age} = 2 \\ 0 & \text{if other} \end{cases} \quad \beta_3 = \begin{cases} 1 & \text{if age} = 3 \\ 0 & \text{if other} \end{cases}$$

Would the linear model itself change, or just the structure of the design matrix (or both)? We'll leave that to you as an 'exercise'. Of course, you should realize by now that while changing the reference age class will change your interpretation of the β 's, it won't change the overall fit of the model to the data.

7.3. multiple 'groups' of marked individuals

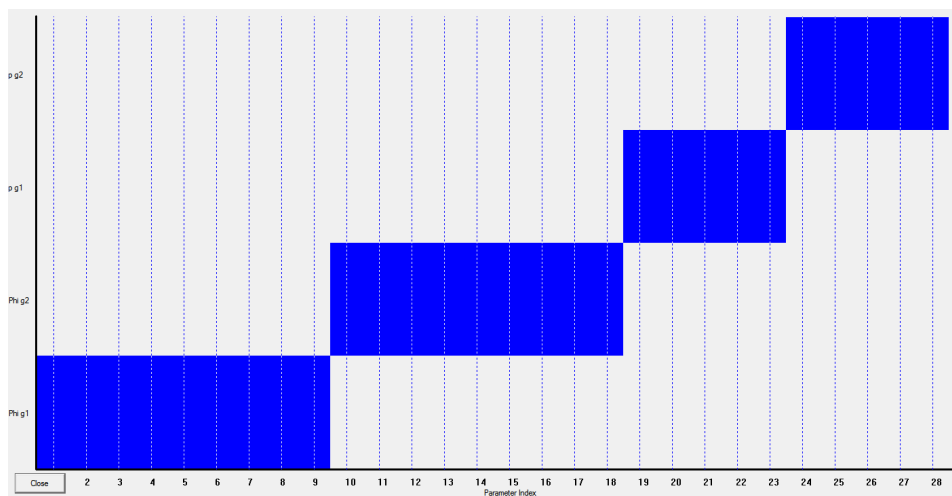
7.3.1. marked as young, multiple marking groups

Suppose you have individuals marked as young (as above), but now, with > 1 groups of individuals (say, you have 2 different breeding populations, that you wish to compare). Superficially, this seems no more difficult than setting up the design matrix for multiple attribute groups, as introduced earlier in Chapter 6. The added complication here though is we also need to model age class. We want to build a linear model that has the following 'factorial structure' (assume we're modeling survival, φ) – let GROUP be the 'marking group' (e.g., the population individuals are marked in), AGE be the age class of the individual at the time it is encountered, and TIME be the interval over which survival is estimated:

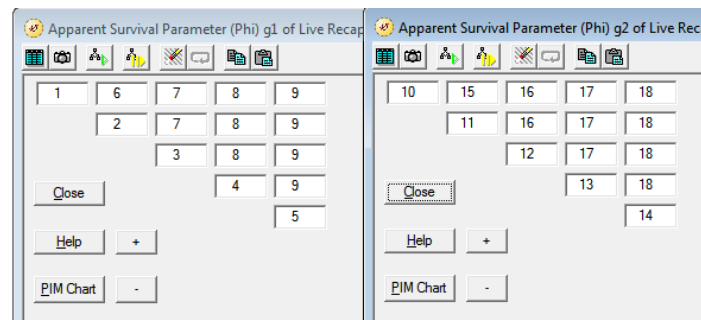
$$\text{logit}(\hat{\varphi}) = \text{GROUP} + \text{AGE} + \text{TIME} + \text{'lots of interactions of \{GROUP, AGE, TIME\}'}$$

As you might have anticipated, the only significant challenge here is working out the interaction terms. We'll illustrate the basic steps using some simulated data in **2grp_2age.inp**: 2 marking groups of individuals all marked as young, with 2 age classes within each group (say, young and adult), and 6 sampling occasions. We want to build / fit a model to these data with group differences, age differences, and full interaction structure, for survival, φ . For p , we'll simply use a full 'GROUP \times TIME' structure. Let's label this model $\{\varphi_{g1:t/t, g2:t/t} p_{g,t}\}$.

Here is the PIM structure for this model:



The indexing for φ_{g1} and φ_{g2} is shown at the top of the next page – note 2 age classes per group, with time dependence for each age class.



Go ahead and fit this model to the data, using the logit link:

Model	AICc	Delta AICc	AICc Weight	No. Par.	Deviance	-2Log(L)
{phi(g1: t1, g2: t1)p(g*t) - PIM - logit}	12382.3079	0.0000	1.00000	26	98.0762	12330.1090

Now, we want to build the design matrix (DM) equivalent. If we look at the PIM chart on the preceding page, we see that the DM should have 28 parameters (columns). Of course, figuring out what goes into those columns is the challenge at hand.

Let's begin by working out the linear model, starting with the 'main effects'. We have 2 groups, so we need $(2 - 1) = 1$ column to code for GROUP. We have 2 age classes, so we also need $(2 - 1) = 1$ column to code for AGE. Then, 6 occasions (5 intervals), so we need $(5 - 1) = 4$ columns to code for TIME:

$$\begin{aligned}\text{logit}(\hat{\phi}) &= \text{GROUP} + \text{AGE} + \text{TIME} + \text{'lots of interactions of \{GROUP, AGE, TIME\}'} \\ &= \beta_1 + \beta_2(\text{GROUP}) + \beta_3(\text{AGE}) + \beta_4(t_1) + \beta_5(t_2) + \beta_6(t_3) + \beta_7(t_4) + \dots\end{aligned}$$

So, 7 columns for the intercept and the main effects. Let's enter these first in the DM. Again make sure you understand how time is coded for the 2 age classes:

B1: phi-int	B2: group	B3: age class	B4: t1	B5: t2	B6: t3	B7: t4
1	1	1	1	0	0	0
1	1	1	0	1	0	0
1	1	1	0	0	1	0
1	1	1	0	0	0	1
1	1	1	0	0	0	0
1	1	0	0	1	0	0
1	1	0	0	0	1	0
1	1	0	0	0	0	1
1	1	0	0	0	0	0
1	0	1	1	0	0	0
1	0	1	0	1	0	0
1	0	1	0	0	1	0
1	0	1	0	0	0	1
1	0	1	0	0	0	0
1	0	0	0	1	0	0
1	0	0	0	0	1	0
1	0	0	0	0	0	1
1	0	0	0	0	0	0

What about the interactions? If you look back at the PIM chart, you'll see that there are 18 survival parameters. So, we'll need to use 18 parameters (columns) in the DM to code for survival. So $(18 - 7) = 11$

columns left for the interactions.

Let's start with the interaction of GROUP and TIME. If you look at the DM at the bottom of the preceding page, you'll see that the first GROUP is coded using the '1' dummy variable, while the second GROUP is coded using the '0'. The interaction of the GROUP column with the TIME columns is simply the product of those columns. So, here are the columns for the interaction of GROUP and TIME:

B1: phi-int	B2: group	B3: age class	B4: t1	B5: t2	B6: t3	B7: t4	B8: g*t1	B9: g*t2	B10: g*t3	B11: g*t4
1	1	1	1	0	0	0	1	0	0	0
1	1	1	0	1	0	0	0	1	0	0
1	1	1	0	0	1	0	0	0	1	0
1	1	1	0	0	0	1	0	0	0	1
1	1	1	0	0	0	0	0	0	0	0
1	1	0	0	1	0	0	0	1	0	0
1	1	0	0	0	1	0	0	0	1	0
1	1	0	0	0	0	1	0	0	0	1
1	1	0	0	0	0	0	0	0	0	0
1	0	1	1	0	0	0	0	0	0	0
1	0	1	0	1	0	0	0	0	0	0
1	0	1	0	0	1	0	0	0	0	0
1	0	1	0	0	0	1	0	0	0	0
1	0	1	0	0	0	0	0	0	0	0
1	0	0	0	1	0	0	0	0	0	0
1	0	0	0	0	1	0	0	0	0	0
1	0	0	0	0	0	1	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0

Next, let's consider AGE and TIME. Again, we need to remember that not all age classes are represented in all time intervals. Specifically, there are no adults in the marked sample in the first interval for either marking group, since all individuals were marked as young. Here is the DM with the interactions of AGE and TIME added:

Design Matrix Specifics													
B1: phi-int	B2: group	B3: age class	B4: t1	B5: t2	B6: t3	B7: t4	B8: g*t1	B9: g*t2	B10: g*t3	B11: g*t4	B12: a*t2	B13: a*t3	B14: a*t4
1	1	1	1	0	0	0	1	0	0	0	0	0	0
1	1	1	0	1	0	0	0	1	0	0	1	0	0
1	1	1	0	0	1	0	0	0	1	0	0	1	0
1	1	1	0	0	0	1	0	0	0	1	0	0	1
1	1	1	0	0	0	0	0	0	0	0	0	0	0
1	1	0	0	1	0	0	0	1	0	0	0	0	0
1	1	0	0	0	1	0	0	0	1	0	0	0	0
1	1	0	0	0	0	1	0	0	0	1	0	0	0
1	1	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	1	0	0	0	0	0	0	0	0	0	0
1	0	1	0	1	0	0	0	0	0	0	1	0	0
1	0	1	0	0	1	0	0	0	0	0	0	1	0
1	0	1	0	0	0	1	0	0	0	0	0	0	1
1	0	1	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	1	0	0	0	0	0	0	0	0	0
1	0	0	0	0	1	0	0	0	0	0	0	0	0
1	0	0	0	0	0	1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0

Note that there is no AGE.TIME interaction column for the first time interval. This is identical to the preceding section, except that here we have > 1 group.

So, 14 out of the 18 columns for survival done. Next, the interaction of GROUP and AGE. Pretty easy – single column. And then, the '3-way' interaction of GROUP, AGE and TIME. Here is the completed DM,

with the '3-way' interaction columns - 18 total columns. Note that there is no '3-way' interaction for the first time interval, since there are no marked adults in the first interval:

B1: phi-intcpt	B2: group	B3: age class	B4: t1	B5: t2	B6: t3	B7: t4	B8: g*t1	B9: g*t2	B10: g*t3	B11: g*t4	B12: a*t2	B13: a*t3	B14: a*t4	B15: group*age	B16: g*a*t2	B17: g*a*t3	B18: g*a*t4	Parm
1	1	1	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1:Phi
1	1	1	0	1	0	0	0	1	0	0	1	0	0	1	1	0	0	2:Phi
1	1	1	0	0	1	0	0	0	1	0	0	1	0	1	0	1	0	3:Phi
1	1	1	0	0	0	1	0	0	0	1	0	0	1	1	0	0	1	4:Phi
1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	5:Phi
1	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	6:Phi
1	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	7:Phi
1	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	8:Phi
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9:Phi
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10:Phi
1	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	11:Phi
1	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	12:Phi
1	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	13:Phi
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14:Phi
1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	15:Phi
1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	16:Phi
1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	17:Phi
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18:Phi

Finally, we might use the following DM structure for the encounter probability, p :

Parm	B19: p-int	B20: p-group	B21: t1	B22: t2	B23: t3	B24: t4	B25: g*t1	B26: g*t2	B27: g*t3	B28: g*t4
1:Phi	0	0	0	0	0	0	0	0	0	0
2:Phi	0	0	0	0	0	0	0	0	0	0
3:Phi	0	0	0	0	0	0	0	0	0	0
4:Phi	0	0	0	0	0	0	0	0	0	0
<several rows edited out...>										
15:Phi	0	0	0	0	0	0	0	0	0	0
16:Phi	0	0	0	0	0	0	0	0	0	0
17:Phi	0	0	0	0	0	0	0	0	0	0
18:Phi	0	0	0	0	0	0	0	0	0	0
19:p	1	1	1	0	0	0	1	0	0	0
20:p	1	1	0	1	0	0	0	1	0	0
21:p	1	1	0	0	1	0	0	0	1	0
22:p	1	1	0	0	0	1	0	0	0	1
23:p	1	1	0	0	0	0	0	0	0	0
24:p	1	0	1	0	0	0	0	0	0	0
25:p	1	0	0	1	0	0	0	0	0	0
26:p	1	0	0	0	1	0	0	0	0	0
27:p	1	0	0	0	0	1	0	0	0	0
28:p	1	0	0	0	0	0	0	0	0	0

So, 28 total columns in the DM, matching the number specified in the PIM.

If we fit this DM-based model to the data, we see we get exactly the same results (meaning, our DM corresponds perfectly to the underlying PIM structure):

Model	AICc	Delta AICc	AICc Weight	No. Par.	Deviance	-2Log(L)
{phi(g1: t1, g2: t2)p(g*t) - PIM - logit}	12382.3079	0.0000	0.50000	26	98.0762	12330.1090
{phi(g1: t1, g2: t2)p(g*t) - DM}	12382.3079	0.0000	0.50000	26	98.0762	12330.1090

If you take your time with the interactions, constructing the DM for > 1 group for 'marked as young' isn't too bad. Alas, things are about to get more complicated...

7.3.2. marked as young + marked as adults

Everything we have covered concerning ‘age models’ has to this point assumed that we were dealing only with individuals all marked as young (or, at the least, at a common age). However, very commonly when we take samples from populations we sample individuals from several age classes, and individually mark any unmarked individuals of all age classes. For example, if our sample contains both newborns/juveniles and adults, we will frequently mark both age classes. Of course, we could choose to just mark the young, but this is often not desirable. Analytically, the question becomes – how to deal with data from both groups of newly marked individuals? In fact, the use of the word ‘group’ was intentional – the solution is simply to treat both types of individuals as different groups with **MARK**.

Previously, we have compared males and females, or controls and treatments, or ‘good’ and ‘poor’ colonies. In this case we simply have ‘marked as young’ versus ‘not marked as young’. In this case, we are testing the hypothesis that survival within an age-class over a given interval doesn’t depend upon the age at which the individual was marked.

How do we do this in **MARK**? In fact, it requires only a slightly more involved application of what we’ve already done. In effect, the analysis becomes one of comparing survival over a given interval between these two ‘groups’ (marked as young and marked as adult). We condition our data based on the age at which the individual was marked: young or adult. For the birds marked as young, we anticipate the possibility of age-specific differences in survival. Let’s assume time-dependence in survival in the first age-class (spanning one year), with time variation in the adult age classes – for this group (i.e., marked as young). Assume 5 occasions. And (as might be an important ‘detail’ which we’ll consider in a moment), assume that the frequencies of encounters of ‘marked as young’ occur before the frequencies of ‘marked as adult’ in the .INP file. Thus, our PIM for birds marked as young would have the following structure:

1	5	6	7
	2	6	7
		3	7
			4

But what about the individuals marked *as* adults? Well, for adults we assume simple time-dependence. We don’t expect any age-specificity in adult survival. So, our PIM file for marked as adults *might* be:

8	9	10	11
	9	10	11
		10	11
			11

We have highlighted the word ‘might’, since the PIMs as we’ve indexed them make an explicit assumption that ‘adult’ survival potentially differs between those individuals originally marked as young, and those individuals originally marked as adults. For example, consider the second interval – for individuals marked as young, the indexing for adult survival (i.e., survival of those individuals who have survived to adulthood) over that interval (parameter index 5) differs from the indexing used for adults over that same interval for individuals originally marked as adults (parameter index 9).

You might be asking yourself ‘why do this – aren’t ‘adults’ simply ‘adults’, regardless of the age they were originally marked?’. The answer is – maybe. Consider the two different groups of individuals – ‘marked as young’, and ‘marked as adult’. What do we know about any ‘adult’ we encounter that was originally ‘marked as young’? Well, we know for sure where it was born, and thus, we might conclude that any ‘adult’ originally ‘marked as young’ is a resident (member) of our sampled population.

On the other hand, consider an ‘adult’ we encounter that was originally ‘marked as adult’. We don’t necessarily know where that individual was born, and thus there might be some (potentially, a lot) of uncertainty concerning whether or not this individual adult is a ‘resident’ or ‘transient’ individual. [We deal explicitly with the analysis of ‘transience’ in section 7.4.] In other words, the ‘marked as adult’ group is potentially a unobservable heterogeneous mix of individuals from different ‘colonies’ (for example), and we might be interested in testing this hypothesis, by first constructing a model where the indexing used for adults differs between ‘marked as young’ and ‘marked as adult’ individuals.

So, given this, and if we start by assuming that the ‘marked as adult’ sample is a potentially heterogeneous mix (such that their survival estimates might differ from the adults encounter from the ‘marked as young’ group), what next? In essence, what we want to determine is whether or not the ‘adult survival’ differs between the 2 groups. In other words, do the estimates for parameters 5 → 7 (adult survival probabilities for birds marked as young) differ significantly from estimates for parameters 9 → 11 (adult survival for individuals marked as adults).

Why is parameter 8 not included in our lists of parameters to compare? If you think about it for a moment, and look at the PIMs we’re starting from, you’ll realize why. Parameter 8 is the probability of a bird marked as an adult at occasion 1 surviving to occasion 2. It cannot be compared with the estimate for parameter 1, which is the survival of an individual marked as a juvenile at occasion 1 to occasion 2. So, we would be comparing ‘apples and oranges’ – a juvenile survival probability with an adult survival probability. By now, you should know exactly how to do this comparison.

In fact, you should realize that there are several ways you could do this – you could modify the PIMs, or modify the design matrix. The most straightforward approach is to modify the PIMs – simply setting the various parameters equal to each other, by using the same index value. Which parameters? $5 = 9, 6 = 10, 7 = 11$.

So, the PIMs would then look like:

1	5	6	7	8	5	6	7
	2	6	7		5	6	7
		3	7			6	7
			4				7

Got it? Make sure you do! Notice that there the main difference is that with individuals marked as both young, and adults, there are both young and adults in the first interval. Thus, the parameter indexing in the PIMs must (in general) reflect this – note that we’ve used 1 for marked as young, and 8 for marked as adults. The actual indexing (numbering) isn’t as important as the structural differences in the PIMs. Since it is very common to work with data containing individuals marked both as young, and as adults, it’s important you truly understand what we’re doing here.

[begin sidebar](#)

‘age before beauty’ (or something clever like that...)

In the preceding we noticed in passing that assumed that the frequencies of encounters ‘marked as young’ occur before the frequencies of ‘marked as adult’ in the .INP file. Under this assumption, we developed the following basic structure (indexing) for the PIMs:

1	5	6	7	8	9	10	11
	2	6	7		9	10	11
		3	7			10	11
			4				11

But, clearly, having 'marked as young' occur before the frequencies of 'marked as adult' in the .INP file is an entirely arbitrary choice. What if instead we flipped it – such that the encounter frequencies of 'marked as adults' occur *before* the frequencies of 'marked as young' in the .INP file? Well, clearly, this would change the PIM structure from the preceding, where the first PIM is 'marked as young' (parameters 1 → 4, juvenile survival, 5 → 7, adult survival), and the second PIM is 'marked as adults' (parameters 8 → 11, adult survival), to the following:

1	2	3	4					5	9	10	11
		2	3	4					6	10	11
			3	4						7	11
				4							8

where the first PIM is now 'marked as adult' (parameters 1 → 4, adult survival), and the second PIM is now the 'marked as young' group (parameters 5 → 8 juvenile survival, 9 → 11, adult survival).

While to the total number of parameters is the same, there may be some conveniences (or additional complications), if you go this route, depending on your question(s), and starting model. For example, if you assume from the outset that adult survival does not differ between 'marked as young and 'marked as adult' individuals, then you could simply set up the PIMs as:

1	2	3	4					5	2	3	4
		2	3	4					6	3	4
			3	4						7	4
				4							8

The DM corresponding to this PIM structure is quite simple – the adults are a contiguous block of 4 rows (for parameters 1 → 4), as are the juveniles (also 4 rows, for parameters 5 → 8).

On the other hand, if you assume from the outset that adult survival does differ between 'marked as young and 'marked as adult' individuals, then the DM corresponding to the 'marked as adults first, marked as young second' PIM structure

1	2	3	4					5	9	10	11
		2	3	4					6	10	11
			3	4						7	11
				4							8

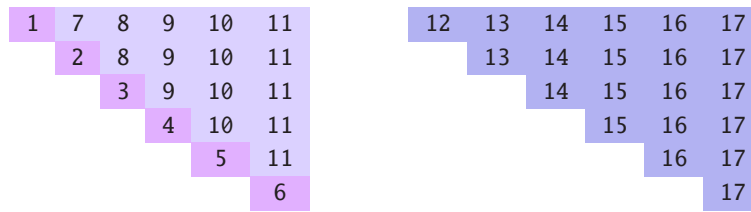
is arguably more complex, since now you have discontinuous blocks of rows for adult parameters (1 → 4, 9 → 11), with juveniles interdigitated between the two (5 → 8). This is tricky (annoying, even) to set up properly. In the end, the best approach to 'ordering the age groups' in your .INP file may depend on choices you make about the starting (general) model.

end sidebar

7.3.3. Marked as young and adult: the 'big ugly' design matrix. . .

Now, for the big test of your understanding of design matrices, and age models. Consider the following problem – again, we focus on a situation where individuals are marked as both young and adults. Assume that our general model is a model with 2 age classes for individuals marked as young, 1 age class for individuals marked as adults. Assume full time-dependence for apparent survival, with no time variation in encounter probability between adults or young, but different encounter probabilities between the two groups. Assume we have 7 occasions. The data for this 'test' are contained in **age_ya.inp**.

Now, we could build our general model using PIMs – here are the apparent survival PIMs, for marked as juveniles, and marked as adults, respectively.



Parameters 1 → 6 represent first-year survival among those marked as young, parameters 7 → 11 corresponds to adult survival for individuals marked as young (remember, for such individuals, there is no adult survival for the first interval, since there are no adults yet!). Parameters 12 → 17 correspond to adult survival for individuals marked as adults, and (finally) parameters 18 and 19 correspond to encounter probability for individuals marked as young, and adult, respectively. Got it? Make sure you do. Then, go ahead and fit this model to the data:

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
(phi(t))p(a) - logit - PIM	13844.2083	0.0000	0.38927	1.0000	19	264.2641

Now, while this is all well and good, and we could build our general model this way using the PIM chart, we realize by now that this is ultimately limiting, since there are some models we can't build using the PIM's. For example, additive models (such as, a model where adult survival and offspring survival for individuals marked as young differed by an additive constant). So, in general, it is recommended you try to build your general model using the design matrix. Your challenge, then, is how to build a design matrix which corresponds to this PIM chart.

Actually, it's not too bad, if you think it through logically. First, we'll start with writing out the linear model for the main effects (as we'll see, there is another, equivalent way to write out the model). We'll let M represent 'age at marking' (one level of grouping), and A represent the 'age of the individual at time of encounter' (note that A is nested within M). As per usual, we'll let 't' represent TIME.

Here is the linear model showing just the main effects:

$$\text{logit}(\hat{\phi}) = \beta_1 + \beta_2(M) + \beta_3(A) + \beta_4(t_1) + \beta_5(t_2) + \beta_6(t_3) + \beta_7(t_4) + \beta_8(t_5) \\ + \text{'a bunch of interactions...'}$$

As indicated, while setting up the main effects is relatively simple, setting up the interactions of the main effects is not quite so straightforward, and will require a bit of thought.

We'll start with a 2-way interaction between age at marking (M) and time (t), because in some senses, it is the easiest. Look again at the PIMs (above). We see clearly that both age at marking groups M (marked as young, marked as adults, respectively) are present in all time intervals.

So, we add the interaction of M and t for all intervals 6 intervals – five additional β parameters ($\beta_9 \rightarrow \beta_{13}$), resulting in a linear model which for the moment has 13 total parameters of the 17 we need (as specified by the PIMs shown at the top of this page):

$$\text{logit}(\hat{\phi}) = \beta_1 + \beta_2(M) + \beta_3(A) + \beta_4(t_1) + \beta_5(t_2) + \beta_6(t_3) + \beta_7(t_4) + \beta_8(t_5) \\ + \beta_9(M.t_1) + \beta_{10}(M.t_2) + \beta_{11}(M.t_3) + \beta_{12}(M.t_4) + \beta_{13}(M.t_5) \\ + \text{'the remaining interactions...'}$$

Now, about the remaining (4) parameters we need? Here, we consider the interaction of age (A) and time (t). Look again at the PIMs (below). Note that there are plausible interactions between age classes (i.e., juveniles, and adults) only for those individuals marked as young (since juveniles don't exist among individuals marked as adults), and only over intervals $2 \rightarrow 6$ (shown shaded in grey):

1	7	8	9	10	11		12	13	14	15	16	17
	2	8	9	10	11			13	14	15	16	17
		3	9	10	11				14	15	16	17
			4	10	11					15	16	17
				5	11						16	17
					6							17

So, we add four β parameters for this interaction between age (A) and time (τ), for the final 5 intervals

$$\begin{aligned} \text{logit}(\hat{\varphi}) = & \beta_1 + \beta_2(\mathbf{M}) + \beta_3(\mathbf{A}) + \beta_4(\mathbf{t}_1) + \beta_5(\mathbf{t}_2) + \beta_6(\mathbf{t}_3) + \beta_7(\mathbf{t}_4) + \beta_8(\mathbf{t}_5) \\ & + \beta_9(\mathbf{M} \cdot \mathbf{t}_1) + \beta_{10}(\mathbf{M} \cdot \mathbf{t}_2) + \beta_{11}(\mathbf{M} \cdot \mathbf{t}_3) + \beta_{12}(\mathbf{M} \cdot \mathbf{t}_4) + \beta_{13}(\mathbf{M} \cdot \mathbf{t}_5) \\ & + \beta_{14}(\mathbf{A} \cdot \mathbf{t}_2) + \beta_{15}(\mathbf{A} \cdot \mathbf{t}_3) + \beta_{16}(\mathbf{A} \cdot \mathbf{t}_4) + \beta_{17}(\mathbf{A} \cdot \mathbf{t}_5). \end{aligned}$$

giving us a final linear model consisting of 17 total parameters – matching the 17 parameters indexed in the PIMs. [Note: some of you might be wondering ‘why no ‘3-way interaction term’ (i.e., why no M.A.t columns?’. Well, beyond the obvious answer that such interactions would require > 17 parameters, which would be inconsistent with the PIMs, the answer is quite simple – there are no young individuals within those marked as adults, so there is no logical 3-way interaction. The same logic applies for ‘why no (A.M) column?’.]

So, we can now proceed to construct the DM corresponding to this linear model (shown at the top of the next page). Pay close attention to the column labels – they’ll help keep track of what is going on. First, the intercept (B1). Then, a column (B2) for the ‘age at marking’ group (M). Then, a column (B3) for ‘age at time of encounter’ (A), followed by the 5 columns (B4 to B8) representing the coding for the 6 time (t) intervals (t1 to t5).

Then the interactions – pay particularly close attention to the coding here.

B1: int	B2: M	B3: A	B4: t1	B5: t2	B6: t3	B7: t4	B8: t5	B9: M:t1	B10: M:t2	B11: M:t3	B12: M:t4	B13: M:t5	B14: A:t2	B15: A:t3	B16: A:t4	B17: A:t5	Pam	B18: p1	B19: p2
1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1:Phi	0	0
1	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	2:Phi	0	0
1	1	1	0	0	1	0	0	0	0	1	0	0	0	1	0	0	3:Phi	0	0
1	1	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	4:Phi	0	0
1	1	1	0	0	0	0	1	0	0	0	0	1	0	0	0	1	5:Phi	0	0
1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6:Phi	0	0
1	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	7:Phi	0	0
1	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	8:Phi	0	0
1	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	9:Phi	0	0
1	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	10:Phi	0	0
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11:Phi	0	0
1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	12:Phi	0	0
1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	13:Phi	0	0
1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	14:Phi	0	0
1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	15:Phi	0	0
1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	16:Phi	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17:Phi	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18:p	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19:p	0	1

Note that for the interaction of M and t , that all of the time coding for both ages for the 'marked as young' group is included. And, for the A and t interactions (nested within the 'marked as young' group), only the plausible interactions for t_2 to t_5 are shown (as is the usual case for 'age models').

Go ahead and run this DM-based model, and add the results to the browser. We see (below) that the model fit is identical to the model we built initially using the PIMs:

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
{ $\phi(t)p(a)$ - logit - PIM}	13844.2083	0.0000	0.38927	1.0000	19	264.2641
{ $\phi(t)p(a)$ - logit - DM - full linear}	13844.2083	0.0000	0.38927	1.0000	19	264.2641

You are urged to study the preceding design matrix carefully. If you understand it, great! This is perhaps one of the more difficult design matrices you can build – but a very common one. If you don't understand it, go back through the preceding page or so, and try again.*

OK, to make sure, a final test (and you thought we were done...). Same data set as above, but now you want to build a particular reduced model, corresponding to the following PIM structure for survival (again we assume 7 occasions in the study).

1	8	9	10	11	12
	2	9	10	11	12
		3	10	11	12
			4	11	12
				5	12
					6

7	8	9	10	11	12
	8	9	10	11	12
		9	10	11	12
			10	11	12
				11	12
					12

The structure of the first PIM (representing individuals marked as young) should be familiar – the diagonal represents the first year after marking, and the off-diagonal represents time-varying survival for these individuals as adults. However, there is a catch – notice that the indexing for the juveniles ends at '6', whereas the adult indexing starts with '8', not '7'. Why?

The reason is explained by looking at the second PIM, representing marked as adults. Here, we use the same indexing (8 → 12) in both PIMs, for adult survival for the 2nd through 6th intervals. In other words, here, we're assuming that adult survival is the same, regardless of whether or not the individuals were marked as young, or marked as adults (whereas earlier, we allowed for adult survival to differ as a function of age at which the individual was marked).

But, what about the use of the index '7' in the first column in the second PIM? We need to come up with a coding reflecting that for those marked as adults, we do in fact have an adult survival estimate over the first interval. But, since there are no adults in that interval for individuals marked as young, we can't constrain the two PIMs to have the same indexing for that interval.

Let's start by building this model using the PIMs themselves (above). At the top of the next page, we've added the results of fitting this model (which we've labeled simply as ' $\phi(t/t)$ - same adult')

* Earlier in this chapter we demonstrated that we can build a DM for models with > 1 group using either a common intercept (where non-intercept β terms reflect the degree to which group estimates differ from the reference value represented by the intercept), or we could use separate intercepts, where each group gets its own intercept. In fact, we can exploit the idea of 'common' vs. 'separate' intercepts for different groups as a 'short-cut' to building models with complex higher-order interaction terms. See the -sidebar- beginning on p. 39.

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
{phi(t)t)p(a) - logit - PIM}	13844.2083	0.0000	0.38927	1.0000	19	264.2641
{phi(t)t)p(a) - logit - DM – full linear}	13844.2083	0.0000	0.38927	1.0000	19	264.2641
{phi(t)t - same adult)p(a) - PIM - logit}	13846.7227	2.5144	0.11073	0.2845	14	276.8261

While we can (and just did) build this model using PIMs, we should be able to construct the identical model by modifying the design matrix corresponding to our most general model.

Look closely again at the design matrix for our general model, shown at the bottom of p. 36. The interaction of 'age of marking' (M, column B2) and time (columns B4 → B8) is shown in columns B9 → B13. The interaction of 'age within age of marking' (A, column B3) and time is shown in columns B14 → B17.

So, how do we build our reduced model, where adult survival is constrained to be the same for individuals marked either as young or adults? Simple, if you realize that our reduced model does **not** have an interaction between time (t) and 'age at marking' (M). In other words, we simply need to eliminate any effect of 'age at marking'.

So, we (1) delete the column corresponding to the 'age at marking' factor, M (i.e., column 2, B2), and then (ii) delete the columns corresponding to the interaction of 'age at marking' (M) and time (i.e., columns B9 → B13).

B1: int	B2: A	B3: t1	B4: t2	B5: t3	B6: t4	B7: t5	B8: At2	B9: At3	B10: At4	B11: At5	Parm	B12: p1	B13: p2
1	1	1	0	0	0	0	0	0	0	0	1:Phi	0	0
1	1	0	1	0	0	0	1	0	0	0	2:Phi	0	0
1	1	0	0	1	0	0	0	1	0	0	3:Phi	0	0
1	1	0	0	0	1	0	0	0	1	0	4:Phi	0	0
1	1	0	0	0	0	1	0	0	0	1	5:Phi	0	0
1	1	0	0	0	0	0	0	0	0	0	6:Phi	0	0
1	0	0	1	0	0	0	0	0	0	0	7:Phi	0	0
1	0	0	0	1	0	0	0	0	0	0	8:Phi	0	0
1	0	0	0	0	1	0	0	0	0	0	9:Phi	0	0
1	0	0	0	0	0	1	0	0	0	0	10:Phi	0	0
1	0	0	0	0	0	0	0	0	0	0	11:Phi	0	0
1	0	1	0	0	0	0	0	0	0	0	12:Phi	0	0
1	0	0	1	0	0	0	0	0	0	0	13:Phi	0	0
1	0	0	0	1	0	0	0	0	0	0	14:Phi	0	0
1	0	0	0	0	1	0	0	0	0	0	15:Phi	0	0
1	0	0	0	0	0	1	0	0	0	0	16:Phi	0	0
1	0	0	0	0	0	0	0	0	0	0	17:Phi	0	0
0	0	0	0	0	0	0	0	0	0	0	18:p	1	0
0	0	0	0	0	0	0	0	0	0	0	19:p	0	1

Are we done? Look carefully – count the columns. We see that the modified DM (above) has 13 total columns. But, according to the PIMs on the preceding page, we need 14 columns! Have we made a mistake?

Well, not exactly, but there is one further step we need to make. Think about the remaining interactions, between age (A) and time (t). Normally, we think of there being one fewer interval for 'adults' in a typical age model. But, now that we've eliminated the differences as a function of 'age at marking' (M), we now have both age classes (i.e., juveniles and adults) in all time intervals. So, we need to add a column for this interaction over the first interval, giving us the needed 14 total columns, as shown below.

Compare this final DM with the preceding one – make sure you see what we’ve done:

B1: int	B2: A	B3: t1	B4: t2	B5: t3	B6: t4	B7: t5	B8: new At1	B9: At2	B10: At3	B11: At4	B12: At5	Param	B13: p1	B14: p2
1	1	1	0	0	0	0	1	0	0	0	0	1:Phi	0	0
1	1	0	1	0	0	0	0	1	0	0	0	2:Phi	0	0
1	1	0	0	1	0	0	0	0	1	0	0	3:Phi	0	0
1	1	0	0	0	1	0	0	0	0	1	0	4:Phi	0	0
1	1	0	0	0	0	1	0	0	0	0	1	5:Phi	0	0
1	1	0	0	0	0	0	0	0	0	0	0	6:Phi	0	0
1	0	0	1	0	0	0	0	0	0	0	0	7:Phi	0	0
1	0	0	0	1	0	0	0	0	0	0	0	8:Phi	0	0
1	0	0	0	0	1	0	0	0	0	0	0	9:Phi	0	0
1	0	0	0	0	0	1	0	0	0	0	0	10:Phi	0	0
1	0	0	0	0	0	0	0	0	0	0	0	11:Phi	0	0
1	0	1	0	0	0	0	0	0	0	0	0	12:Phi	0	0
1	0	0	1	0	0	0	0	0	0	0	0	13:Phi	0	0
1	0	0	0	1	0	0	0	0	0	0	0	14:Phi	0	0
1	0	0	0	0	1	0	0	0	0	0	0	15:Phi	0	0
1	0	0	0	0	0	1	0	0	0	0	0	16:Phi	0	0
1	0	0	0	0	0	0	0	0	0	0	0	17:Phi	0	0
0	0	0	0	0	0	0	0	0	0	0	0	18:p	1	0
0	0	0	0	0	0	0	0	0	0	0	0	19:p	0	1

Go ahead and fit this model using the modified DM to the data. You should see (browser, below) that it produces the identical fit as did the equivalent model we constructed using the PIMs:

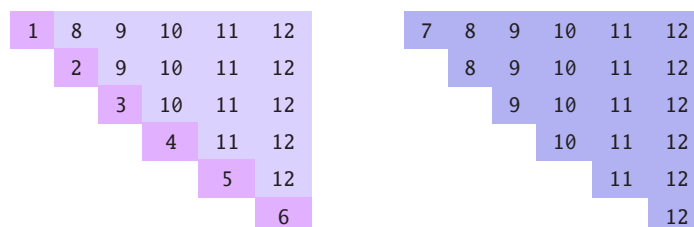
Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
$\phi(t)p(a) - \text{logit} - \text{PIM}$	13844.2083	0.0000	0.38927	1.0000	19	264.2641
$\phi(t)p(a) - \text{logit} - \text{DM} - \text{full linear}$	13844.2083	0.0000	0.38927	1.0000	19	264.2641
$\phi(t - \text{same adult})p(a) - \text{PIM} - \text{logit}$	13846.7227	2.5144	0.11073	0.2845	14	276.8261
$\phi(t - \text{same adult})p(a) - \text{logit} - \text{modified DM}$	13846.7227	2.5144	0.11073	0.2845	14	276.8261

```
begin sidebar
```

a clever (but limited) trick for building 'big ugly DM with interactions...'

The preceding several pages have illustrated the ‘challenges’ of building complex age’ models that involve higher-level interaction terms. However, as suggested at various points, there is an approach for constructing the DM for the general model with ‘full interactions’ that is quite straightforward to implement – if you are not interested in modifying this DM to build additive models. This approach involves using a separate intercept for each group.

We will demonstrate this by re-visiting the preceding analysis of data from individuals marked as both young and adults (section 7.3.3) – this analysis involved the most complex DM in this chapter, with multiple interaction ‘columns’. Here again is the PIM structure for our general model



and the corresponding DM with a single, common intercept:

B1: int	B2: M	B3: A	B4: t1	B5: t2	B6: t3	B7: t4	B8: t5	B9: Mt1	B10: Mt2	B11: Mt3	B12: Mt4	B13: Mt5	B14: At2	B15: At3	B16: At4	B17: At5	Parm	B18: p1	B19: p2
1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1:Phi	0	0
1	1	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	2:Phi	0	0
1	1	1	0	0	1	0	0	0	0	1	0	0	0	1	0	0	3:Phi	0	0
1	1	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	4:Phi	0	0
1	1	1	0	0	0	0	1	0	0	0	0	1	0	0	0	1	5:Phi	0	0
1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6:Phi	0	0
1	1	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	7:Phi	0	0
1	1	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	8:Phi	0	0
1	1	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	9:Phi	0	0
1	1	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	10:Phi	0	0
1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11:Phi	0	0
1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	12:Phi	0	0
1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	13:Phi	0	0
1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	14:Phi	0	0
1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	15:Phi	0	0
1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	16:Phi	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17:Phi	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18:p	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19:p	0	1

From the PIM structure, we see (in effect) 3 groups: juveniles marked as juveniles, adults marked as juveniles, and adults marked as adults. So, if we use a separate intercept for each group, this means 3 intercepts. Here is the DM reflecting this:

B1 int-juv (MJ)	B2 t1	B3 t2	B4 t3	B5 t3	B6 t5	B7 int-adt (MJ)	B8 t2	B9 t3	B10 t3	B11 t5	B12 int-adt (MA)	B13 t1	B14 t2	B15 t3	B16 t4	B17 t5	Parm	B18 p
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1:Phi	0
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2:Phi	0
1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3:Phi	0
1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	4:Phi	0
1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	5:Phi	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6:Phi	0
0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	7:Phi	0
0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	8:Phi	0
0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	9:Phi	0
0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	10:Phi	0
0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	11:Phi	0
0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	12:Phi	0
0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	13:Phi	0
0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	14:Phi	0
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	15:Phi	0
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	16:Phi	0
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	17:Phi	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18:p	1

Columns 1 → 6 model juveniles marked as juveniles, 7 → 11 model adults marked as juveniles, and 12 → 17 model adults marked as adults. If you build and fit this DM to the data used in section 7.3.3, you'll see (by looking at the model deviance or $-2\ln(\mathcal{L})$) the fits are identical.

Using separate intercepts to build the 'full interaction' model is straightforward, and much simpler (mechanically, and perhaps conceptually) than using the single 'common intercept' approach we have generally used. However, a DM constructed using separate intercepts for different groups cannot be easily modified to construct models nested within the general model (e.g., additive models).

end sidebar

7.4. 'Time since marking' – when an age model is NOT an 'age' model

In the preamble to this chapter we noted that 'age' models are generally motivated by the fact that we expect that individuals of different ages might have different survival or recapture probabilities. Also recall that we noted that within a cohort, age and time were synonymous (and thus, age and time effects could not be separately estimated, at least within a cohort). In fact, the age of an animal is always defined as the time since the animal was marked. If the animal was marked as a newborn, then the time since marking is equivalent to true chronological age. But, what if the animal was marked as an adult – say, as a 3 year old? If the year of marking is year (t), and it is currently year ($t + 5$), then the animal is chronologically 8 years old, but it is 5 years old, relative to the age at which it was marked.

Why do we care? We care because in many instances, we do not expect to detect true age differences, but rather, differences among individuals as a function of the relative time since marking. Recall that in a true age model, we anticipate that the survival (for example) of newborns might be different (typically lower) than the survival of adults. Structurally, this is equivalent to saying that we anticipate that the survival the year after marking (which in this case refers to the first year of life) will differ from the survival in subsequent years of life (i.e., among those individuals surviving the first year of life, we expect these surviving individuals – now adults – to have different survival in all subsequent years). So, in essence, we are considering variation in survival (in this example) as a function of 'time since marking' (which we will refer to as '*TSM models*'). So, hereafter, we distinguish between 'true' age models (where we are interested in true, chronological age effects) from TSM models, where we're not interested in age, *per se*, but the time elapsed since the animal was marked (i.e., relative age).

A reasonable question at this stage is, 'Why do we need to consider TSM models at all?'. As mentioned, there are a number of common situations where it is the time since marking which influences various parameters. We'll introduce one fairly well-studied example to illustrate the use of TSM models: the presence of *transient* individuals in marked samples. The consideration of transience in marked samples begins with the seminal paper by Pradel and colleagues (1997). We follow Pradel *et al.* and define a transient individual as 'an individual that is marked, released, and which then permanently emigrates from the sample, such that it is no longer available for encounter in the future'.

In other words, a transient is an individual that is seen once (the marking event), then never seen again, because it has emigrated from the population. Recall that in a classic live encounter study, we cannot separate true death from permanent emigration (they are confounded). Thus, a transient individual, which by definition permanently emigrates from the population, will appear to have 'died'. (We realize that some of the more biologically-minded readers will at this point be wondering about 'temporary emigration' – the case where an individual leaves the sample, but not permanently. Temporary emigration is dealt with in a later chapter (15) on the 'robust design'.

We differentiate transient individuals (as defined previously) from resident individuals, which are those individuals that are marked, released, and stay in the sample. Conditional on remaining alive, these resident individuals thus have the potential to be encountered again, after the initial marking event. The presence of both resident and transient individuals in the sample is a clear violation of one of the assumptions of the classical CJS model – namely, that all individuals have the same probability of subsequent encounter. The presence of residents and transients violates this assumption, since transients (which have an encounter probability after marking of 0) do not have the same probability of encounter as do residents. Thus, the population is said to show heterogeneity among individuals, in one or more parameters. Earlier in the chapter, we allowed for differences among individuals in terms of true chronological age. But, how do we account for differences among residents and transients in the present example?

The clue to answering this is to recall that a transient is an individual which permanently emigrates after marking. If we start by assuming that this emigration 'event' occurs during the interval after marking, we might start to see a connection with TSM models. If not, the following numerical example should make this clear. Suppose that we have a population of some bird species, where the annual survival of residents is a constant, $\varphi_{Re} = 0.8$. Suppose that the encounter probability for this population is also constant, and equal to 1.0 (i.e., there is perfect, complete registration of all living resident individuals). Now, what about transients? Well, if a transient permanently emigrates immediately after marking, then its apparent survival probability is $\varphi_{Tr} = 0$.

Now, imagine we start with a sample of 100 individuals from some population. Assume that they are all adults of the same age at the time of marking (thus, no heterogeneity in true age). Suppose that 50% of them are resident, and 50% are transient. Of course, in the 'real world' we wouldn't know this, but for the moment, we'll assume that we do. So, $R_1 = 100$. Given that $p = 1$, then how many individuals from this release cohort would we expect to see at occasion 2? Fairly easy calculation, if you remember that the apparent survival probability for residents is 0.8, and for transients is 0. Since 50 individuals in R_1 are residents, then we expect that 80% of them (40/50) and be encountered at occasion 2. In contrast, of the 50 transient individuals in R_1 , all of them will emigrate, and thus no transients will be encountered at occasion 2.

Thus, what we would 'see' in our data is $R_1 = 100$, and the number encountered at occasion 2 is 40. If we didn't know the starting, true proportions of residents and transients in the population, then our estimate of survival for the 100 marked and released individuals would be $(40/100) = 0.4$. Obviously, this is not a good estimate of true resident survival probability. It is biased low – we estimate an apparent survival probability of 0.4, but the true survival probability is 0.8! Clearly, our estimate is negatively biased.

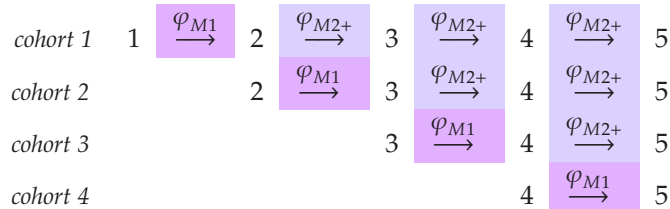
What about the interval between occasion 2, and occasion 3. Recall that at occasion 1, our release sample R_1 consisted of both residents and transients. However, by occasion 2, our sample consists entirely of residents (since by definition the transients all emigrated during the interval between occasion 1 and 2). Thus, over the interval from occasion 2 to occasion 3, we would expect the estimated survival to reflect that for residents (i.e., $\varphi_{Re} = 0.8$), since there are only residents released at occasion 2. And so on for the interval from occasion 3 to occasion 4, occasion 4 to occasion 5. . .

Now, what does this have to do with TSM models? Well, from the preceding, it is clear that during the interval after marking, the marked sample is a mixture of residents and transients – let's call the apparent survival probability estimated over this interval φ_{M1} , where 'M1' refers to the first year after marking – the M indicating 'marking'. Let φ_{M2+} be the survival rate estimated during the intervals after the first interval following marking, where 'M2+' refers to years 2 and over after marking).

Thus, for a single release cohort, we could write:



Now, suppose that we release newly marked individuals at each occasion – if we have 5 total occasions, we could write



Which, in **MARK** (PIM) format, is equivalent to

1	2	2	2
	1	2	2
		1	2
			1

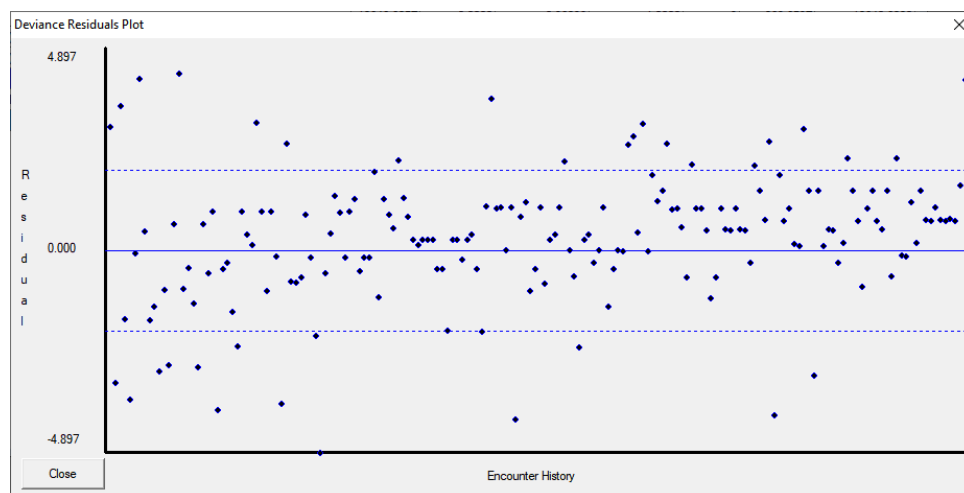
Look familiar? It should – it's identical to the PIM for a 2 age-class model, with no time variation in either age class! Here, though, we are **not** dealing with true 'age' differences, but an effect of heterogeneity in the sample. And yet, the PIM is identical to what we might use for an 'age' model. And, therein lies the potential for confusion – if you use a PIM with 'age' structure to analyze a data set that doesn't have real age effects (say, a data set where everyone is marked as an adult), you are bound to confuse someone who may not know that an 'age' model refers to a particular model structure, and not necessarily to true age. So, we recommend use of 'TSM models' as the generic name for such models – if you really are dealing with 'true' age effects (such as might be the case when you have individuals marked as young), then you probably should call them 'age models'.

To reinforce the utility of TSM models, let's consider an example of some simulated live encounter data which contains both resident and transient animals (**transient.inp**). At each of 8 occasions, we simulated the release of 500 newly marked individuals. To keep things simple, we assumed that at each occasion, the sample consisted of the same proportion of residents to transients (70% residents, 30% transients). We assumed that survival of residents was constant, 0.8, and that the encounter probability of residents (conditional on surviving) was 0.7.

To reinforce the typical steps in an analysis, we'll start with a GOF test of a general model to these data. Now, you might have reason to suspect transients in this population, enough that you might make an *a priori* decision to fit a general model with structure which accounts for transients (the TSM models we're introducing here, for example). But, for the moment, let's 'pretend' we know nothing about these data – under such circumstances, we typically would start with a standard 'time-dependent' model. We'll do so here – we'll fit model $\{\varphi_t p.\}$ to the data.

Go ahead and run this model in **MARK**. If we run a median \hat{c} GOF test, we find that the estimated \hat{c} is ~ 1.9 . While this is not unreasonably large (i.e., it does fall below our rule-of-thumb $\hat{c} \leq 3$), it does give us some indication of lack-of-fit of this model to the data.

Looking at the model residuals (below) seems to confirm this:



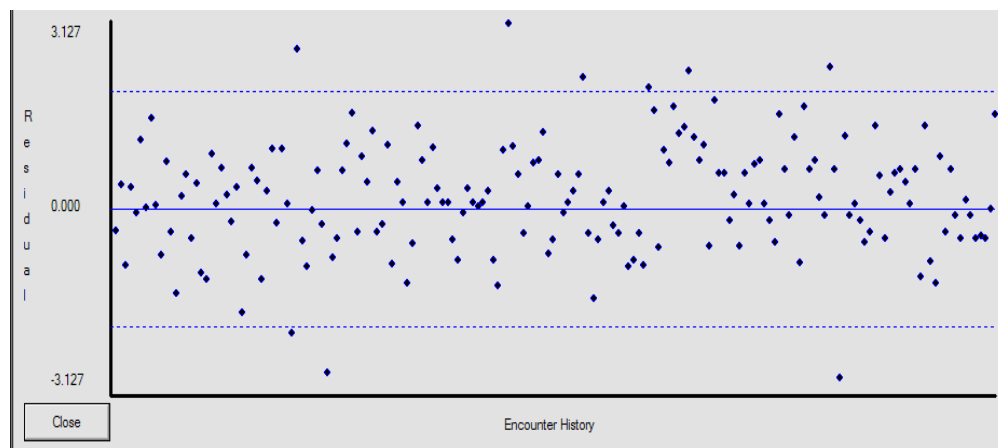
If the model was a 'good-fitting model', we would expect that there would be no 'trend' (non-randomness) in the pattern of the residuals - half of the residuals should be above the 0.000 line, and half should be below. Further, if there is no extra-binomial variation, then most of these residuals should be fairly close to the 0.0000 line (between the two horizontal dashed lines in the figure). This is not what we see – there is a clear asymmetry of distribution of the points around the 0.000 line (in this case, more of the residuals are above the 0.000 than are below it), and a fair number of the residuals are well above/below the dashed lines. Given this, together with the estimate of $\hat{c} = 1.9$, would suggest that we need to find a better general starting model.

So, let's try fitting a TSM model – we might 'suspect' this is likely to be a better model structure for these data. We'll fit the TSM model with 2 M -classes, constant over time in each class (call this model $\{\phi_{M2-.}, p.\}$). Here are the results:

$\{\phi_{M2-.}, p.\}$	12646.6257	0.0000	1.00000	1.0000	3	263.6537
$\{\phi(t)p.\}$	12921.1302	274.5045	0.00000	0.0000	8	528.1391

The model with 2 M -classes is the best model – it has 100% of the weight! This isn't surprising since it is the 'true' model for the simulated data. The point of note, however, is that a model with 'age' in the structure fits these data better than a simple time-dependent model, even though there are no true age effects in the data.

We can confirm this by looking at the estimate of \hat{c} for this model (~ 1.01), and by looking at the residual plot, below:



We see that in this case, the residuals are distributed evenly above/below the 0.000 line, indicative of pretty good fit of the model to the data.

Further, if we look at the estimates from the most parsimonious model

Real Function Parameters of $\{\phi_{M2-.}, p.\}$				
Parameter	Estimate	Standard Error	95% Confidence Lower	95% Confidence Upper
1:Phi	0.5612224	0.0098753	0.5417834	0.5804754
2:Phi	0.8067660	0.0074555	0.7917320	0.8209602
3:p	0.7016322	0.0081143	0.6854887	0.7172889

we see that the estimate of survival of the first M -class (i.e., in the interval following marking) is negatively biased, 0.5612 (as expected), whereas the estimate of survival for subsequent intervals, 0.8067, is very close to the ‘true’ underlying parameter value 0.8.

Recall that after the first interval, the marked sample for a given cohort consists entirely of residents, who will survive at the resident survival probability (0.8, in this case). The estimate of p is not influenced by transient individuals (since they are all out of the sample by the second occasion) – remember that p is estimated conditional on surviving and remaining in the sample. So, clearly, transient individuals do not influence the estimate of p .

So, by fitting a TSM model to these data, we were able to derive an unbiased estimate of resident survival probability – we accomplish this by accounting for differences in survival as a function of time since marking-in this case, time since marking corresponds to changes not in age, but in the heterogeneity of the marked individuals in the sample.

begin sidebar

proportion of transients in newly marked sample

For completeness, we note that from the estimates of survival for the 2 M -class, we can derive an estimate of the proportion of *residents* in the ‘newly marked’ sample of individuals for a given interval by taking $\hat{\phi}_{T+M}$ (i.e., estimated survival over the a particular interval where *both* residents and transients are in the marked sample) and dividing it by $\hat{\phi}_{Re}$ (i.e., the estimate of survival over the same interval when only residents are in the sample). In other words, the estimates along the diagonal divided by the off-diagonal estimates, for a given interval – i.e., within a column of the PIM.

In this case, where the underlying probabilities and proportions of residents and transients in the newly marked sample were constant over time (since we simulated them that way), the proportion of *residents* would be given as $0.561/0.807 = 0.695 \sim 70\%$, which was the true value used in the simulation. Clearly, $1 - (0.561/0.807)$ gives the proportion of *transients* in the marked sample (*note*: in the sample, **not** necessarily in the population).

Why does this work? Simple algebra. Let the total number of birds released at occasion 1 for a given cohort be $(N_{Tr} + N_{Re})$, where N_{Tr} = number of transients in the sample, and N_{Re} = number of residents in the sample. Since only residents will be encountered at occasion 2 (conditional on surviving), then the number of individuals expected alive at occasion 2 is $\phi_{Re} N_{Re}$ (where ϕ_{Re} is the survival probability of residents). Thus, the apparent survival probability over the first interval is $\phi_{Re} N_{Re} / (N_{Re} + N_{Tr})$. From occasion 2 to occasion 3, within the same cohort, the population consists entirely of residents. The survival probability of residents is ϕ_{Re} . Thus, the survival in the first interval divided by the survival after the first interval is $[\phi_{Re} N_{Re} / (N_{Re} + N_{Tr})] / \phi_{Re} = N_{Re} / (N_{Re} + N_{Tr})$, which is the proportion of residents in the sample! For details, see the Pradel *et al.* (1997) paper.

To get an estimate of the SE for the proportion of residents (or transients), we could use the Delta method (described in Appendix B).

end sidebar

7.4.1. Age, transience and the DM – a complex example

We end this section with consideration of a more complex example. The purpose of this example is to reinforce our understanding of building design matrices for TSM models. Such models are arguably among the most complicated you are likely to work with in practice, and if you are able to work through this example, you can be fairly confident that you’re likely to be able to handle even complex problems. The basic idea(s) are very similar to what covered earlier, in section 7.3.1.

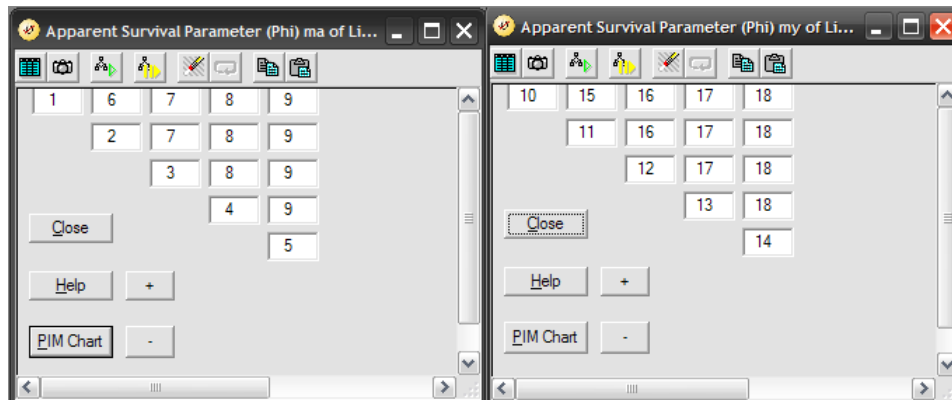
Here we imagine a 6 occasion study where have individuals marked as young, and adults. Amongst those marked as young, we anticipate temporal variation in survival over the first year (i.e., within the

first TSM class), but constant among individuals surviving to become adults. Among those marked as adults, we suspect that the first TSM class consists of some temporally variable mixture of resident and transient individuals (defined as in the preceding example). After the first year after marking (i.e., after the first TSM class), we anticipate adult survival is constant over time.

However, we have some reason to suspect that survival of adults may differ as a function of the age at which individuals were initially marked and released. We expect that encounter probability is constant over time, with no TSM structure, and does not differ as a function of age at marking. The data for this example problem are contained in **age-transients-full.inp**. Note that in this input file, the encounter frequencies for the ‘marked as adult’ group come first, followed by the encounter frequencies for the ‘marked as young’ group).

Start a new project, remembering that you have 2 attribute groups, which we’ll label ‘ma’ and ‘my’ (corresponding to ‘marked as adult’ and ‘marked as young’, respectively). Since this is a fairly complex problem, we’ll start by constructing our general model using PIMs. For our general model, we’ll assume full time dependence throughout. We’ll label as $\{\varphi_{ma:t/t\ my:t/t,\Delta adt\ p.}\}$, where ma refers to ‘marked as adults’, my refers to ‘marked as young’, and Δadt refers to ‘adult survival differing as a function of age of marking’.* The PIMs corresponding to apparent survival φ for ‘marked as adults’ and ‘marked as young’ are shown at the top of the next page.

For the ‘marked as adults’ group (left-hand PIM), parameters 1 \rightarrow 5 correspond to the first TSM class (which contains a mixture of transients and residents), while parameters 6 \rightarrow 9 correspond to ‘adult’ (ie., resident) survival. For the ‘marked as young’ group (right-hand PIM), parameters 10 \rightarrow 14 correspond to the first TSM class (i.e., age 0 \rightarrow 1), while parameters 15 \rightarrow 18 correspond to the adults. With constant and equal encounter probability for both groups, we have 19 total parameters in our model.



Run this model, and add the results to the browser:

Results Browser: Live Recaptures (CJS)						
Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
$\phi(ma:t/t\ my:t/t) - \text{different adult} - \text{PIM}$	2332.4441	0.0000	1.00000	1.0000	19	113.5508

* Coming up with a ‘naming convention’ for TSM models can prove somewhat of a challenge.

Note that we add the word ‘PIM’ to the model name to indicate that this version of the general model was constructed using PIMs. We will use this model to ‘confirm’ that the we have constructed the equivalent DM-based model correctly.

Let’s start the construction of the DM corresponding to the parameter structure specified by these PIMs by writing out the first few terms in our linear model, corresponding to the ‘main effects’:

$$\varphi = \text{intcpt} + \text{M} + \text{A} + \text{T},$$

where M refers to the ‘age of marking’, A refers to ‘age’ (or TSM class) within age of marking group), and T refers to time. Since there are 2 age of marking groups, then we need one column in the DM to code for M. Since there are 2 age (TSM) classes within each age of marking group, we need one column in the DM to code for A. And finally, since there are 6 occasions, there are 5 intervals, so we need four columns in the DM to code for T. So, our linear model so far is

$$\begin{aligned}\varphi &= \text{intcpt} + \text{M} + \text{A} + \text{T} \\ &= \beta_1 + \beta_2(\text{M}) + \beta_3(\text{A}) + \beta_4(\text{T}_1) + \beta_5(\text{T}_2) + \beta_6(\text{T}_3) + \beta_7(\text{T}_4)\end{aligned}$$

Let’s start our DM by coding these 7 terms (shown below):

B1 intcpt	B2 M	B3 A	B4 T1	B5 T2	B6 T3	B7 T4	Parm
1	1	1	1	0	0	0	1:Phi
1	1	1	0	1	0	0	2:Phi
1	1	1	0	0	1	0	3:Phi
1	1	1	0	0	0	1	4:Phi
1	1	1	0	0	0	0	5:Phi
1	1	0	0	1	0	0	6:Phi
1	1	0	0	0	1	0	7:Phi
1	1	0	0	0	0	1	8:Phi
1	1	0	0	0	0	0	9:Phi
1	0	1	1	0	0	0	10:Phi
1	0	1	0	1	0	0	11:Phi
1	0	1	0	0	1	0	12:Phi
1	0	1	0	0	0	1	13:Phi
1	0	1	0	0	0	0	14:Phi
1	0	0	0	1	0	0	15:Phi
1	0	0	0	0	1	0	16:Phi
1	0	0	0	0	0	1	17:Phi
1	0	0	0	0	0	0	18:Phi

After the single column for the intercept, we code the M effect: we use 1’s to indicate ‘marked as adult’, and 0’s to indicate ‘marked as young’. This is followed by the column coding for the A effect – here we use 1’s to code for the first TSM class within each age of marking group, and 0’s to code for the second TSM class. Recall that for ‘marked as adults’ the first TSM class reflects a mixture of residents and transients, whereas for the ‘marked as young’ group, the first TSM class are juveniles. Finally, 4 columns to code for each of the time intervals (here we use the final time interval as the reference interval). Note again that the second TSM class is not represented in the first time interval (and thus the time coding for the second TSM class begins with the second interval).

Now, for the interactions of the main effects, which is typically the largest challenge. The key is in remembering that coded interactions must be ‘plausible’. So far, we have 7 parameters for φ in our DM, meaning, we have $(18 - 7) = 11$ parameters remaining for the interaction terms. Which in turn means

that if we have more or less than 11 columns for our interactions, we've made a mistake somewhere.

What are the 'possible' interactions? For a model with 3 main effects (M, A and T), there are 3 possible two-way interactions (M.A, M.T, and A.T), and 1 possible three-way interaction (M.A.T). Remember, if any of the two-way interactions aren't plausible, then neither is the three-way interaction (since the three-way interaction consists of the product of a given two-way interaction and the remaining effect).

So, in theory, our linear model could look like

$$\begin{aligned}\varphi = & \text{intcpt} + M + A + T \\ & + M.A + M.T + A.T \\ & + M.A.T\end{aligned}$$

So, which of the 'possible' interactions are 'plausible'? Start with the (M.A) interaction. We see from the DM (above) that there is 'symmetry' between the two groups ('marked as adult' and 'marked as young') in the TSM structure. So, there is a plausible (M.A) interaction term. Since the (M.A) interaction is the product of two single columns, then the (M.A) interaction is coded as a single column. So, that would give us $(7 + 1) = 8$ columns; 10 remaining.

What about the (M.T) interaction? Again, we see that the time columns are symmetrical between the two age of marking groups. In other words, the time coding for 'marked as adults' is perfectly replicated for the 'marked as young' group. So, we have a full (M.T) interaction (1 column for M with 4 columns for time, yield 4 columns for the M.T interaction term). So, $(8 + 4) = 12$ columns coded so far; 8 left.

Next, the (A.T) interaction. Here, we simply need to remember that an (A.T) interaction is not plausible for all time steps. Since there are no individuals representing the second TSM class in the first time interval for either the 'marked as adult' or 'marked as young' groups, then (A.T) interactions are plausible only for time intervals $2 \rightarrow 5$. So, 3 columns for the (A.T) interaction. So, to this point, $(12 + 3) = 15$ columns; so, 3 left.

Finally, the three-way interaction, (M.A.T). All we need to do at this point is take the (A.T) interaction, and multiply through by the M term. We know that we have 3 columns left to code for the three-way interaction. Fortunately 1 (for M) times 3 (for the (A.T) interaction columns) = $3!$

The completed design matrix for our general model is shown below:

Design Matrix Specification (B = Beta)																			
B1 intcpt	B2 M	B3 A	B4 T1	B5 T2	B6 T3	B7 T4	B8 M.A	Pam	B9 M.T1	B10 M.T2	B11 M.T3	B12 M.T4	B13 A.T2	B14 A.T3	B15 A.T4	B16 M.A.T2	B17 M.A.T3	B18 M.A.T4	B19 p
1	1	1	1	0	0	0	1	1:Phi	1	0	0	0	0	0	0	0	0	0	0
1	1	1	0	1	0	0	1	2:Phi	0	1	0	0	1	0	0	1	0	0	0
1	1	1	0	0	1	0	1	3:Phi	0	0	1	0	0	1	0	0	1	0	0
1	1	1	0	0	0	1	1	4:Phi	0	0	0	1	0	0	1	0	0	1	0
1	1	1	0	0	0	0	1	5:Phi	0	0	0	0	0	0	0	0	0	0	0
1	1	0	0	1	0	0	0	6:Phi	0	1	0	0	0	0	0	0	0	0	0
1	1	0	0	0	1	0	0	7:Phi	0	0	1	0	0	0	0	0	0	0	0
1	1	0	0	0	0	1	0	8:Phi	0	0	0	1	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	9:Phi	0	0	0	0	0	0	0	0	0	0	0
1	0	1	1	0	0	0	0	10:Phi	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	1	0	0	0	11:Phi	0	0	0	0	1	0	0	0	0	0	0
1	0	1	0	0	1	0	0	12:Phi	0	0	0	0	0	1	0	0	0	0	0
1	0	1	0	0	0	1	0	13:Phi	0	0	0	0	0	0	1	0	0	0	0
1	0	1	0	0	0	0	0	14:Phi	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	1	0	0	0	15:Phi	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	1	0	0	16:Phi	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	1	0	17:Phi	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	18:Phi	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	19:p	0	0	0	0	0	0	0	0	0	0	1

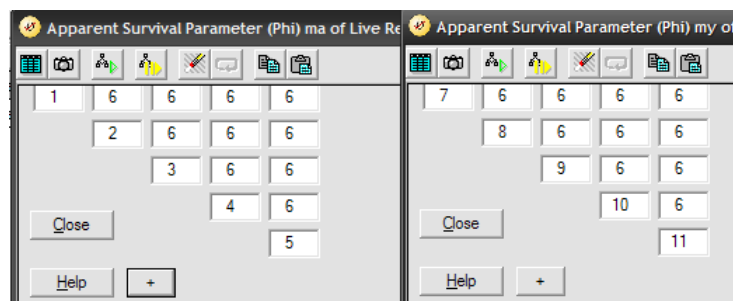
To confirm that we've have correctly constructed the DM, run the model, making sure to change the model name to indicate that we're using the DM, and add the results to the browser:

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance
{ma: t/t my: t/t - different adult - PIM}	2332.4441	0.0000	0.50000	1.0000	19	113.5508
{ma: t/t my: t/t - different adult - DM}	2332.4441	0.0000	0.50000	1.0000	19	113.5508

Notice that the model deviances between the model built with PIMs, and the one we just fit using the DM approach, are identical. This indicates that our DM is correct.

A final test of our understanding. Suppose we want to build a reduced parameter model, where survival varies temporally for the first TSM class for both 'marked as adult' and 'marked as young' groups, but with no time variation in the second TSM class, and the same value between marking groups (in other words, time-invariant survival of the second TSM class that does not vary as a function of the age of marking).

The PIM structure for this model would look something like that shown below:



But, recall that we'd like to build these models by modifying the DM of our general model. Here is the modified DM which corresponds to this reduced parameter model:

Design Matrix Specification (B = Beta)												
B1: Int	B2: M	B3: A	B4: T1	B5: T2	B6: T3	Parm	B7: T4	B8: M.T1	B9: M.T2	B10: M.T3	B11: M.T4	B12: p
1	1	1	1	0	0	1:Phi	0	1	0	0	0	0
1	1	1	0	1	0	2:Phi	0	0	1	0	0	0
1	1	1	0	0	1	3:Phi	0	0	0	1	0	0
1	1	1	0	0	0	4:Phi	1	0	0	0	1	0
1	1	1	0	0	0	5:Phi	0	0	0	0	0	0
1	0	0	0	0	0	6:Phi	0	0	0	0	0	0
1	0	0	0	0	0	7:Phi	0	0	0	0	0	0
1	0	0	0	0	0	8:Phi	0	0	0	0	0	0
1	0	0	0	0	0	9:Phi	0	0	0	0	0	0
1	0	1	1	0	0	10:Phi	0	0	0	0	0	0
1	0	1	0	1	0	11:Phi	0	0	0	0	0	0
1	0	1	0	0	1	12:Phi	0	0	0	0	0	0
1	0	1	0	0	0	13:Phi	1	0	0	0	0	0
1	0	1	0	0	0	14:Phi	0	0	0	0	0	0
1	0	0	0	0	0	15:Phi	0	0	0	0	0	0
1	0	0	0	0	0	16:Phi	0	0	0	0	0	0
1	0	0	0	0	0	17:Phi	0	0	0	0	0	0
1	0	0	0	0	0	18:Phi	0	0	0	0	0	0
0	0	0	0	0	0	19:p	0	0	0	0	0	1

Pay particular attention to the way we've modified the M coding: since we've eliminated any effect of age of marking on survival for the second TSM class, we eliminate the 1's coding for the second TSM class for the 'marked as adults' group. The rest of the DM should be fairly self-explanatory.

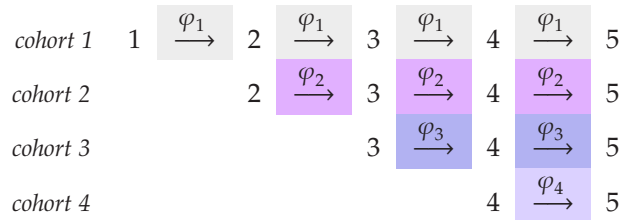
7.5. Age/TSM models and GOF

If you remember back to Chapter 5, you may recall that there are various ways to test for GOF for a given model. Looking at deviance residual plots was described above. Note that because age/TSM models have more parameters than standard time-dependent CJS models, then in fact age/TSM models are more 'general' than CJS models, and as such, should be models for which you assess fit (if such models are included in your candidate model set). You can use either **RELEASE** (for model with 2 'classes', by adding some of the individual component tests – discussed in Chapter 5), the bootstrap, median- \hat{c} or Fletcher- \hat{c} methods, or program **U-CARE**, which has a large number of component tests specific for age/TSM models (including transience testing, for example).

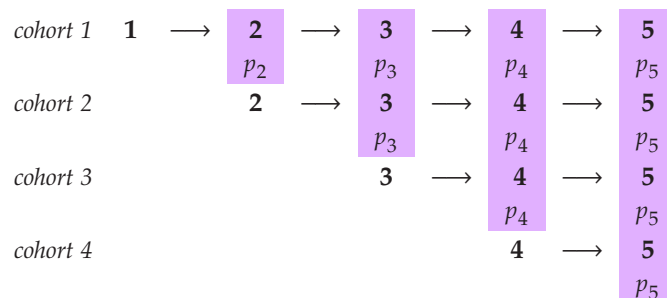
7.6. Cohort models

The time-dependent CJS model assumes that neither survival nor recapture differ among release cohorts. Is this a reasonable assumption? Suppose, for example, that animals newly marked on occasion 3 were present in the population on occasion 1, but were simply 'missed' from the marking sample. Perhaps there is a 'reason' why these animals were missed – and this reason might influence subsequent survival or (perhaps more likely) recapture probability. Another common example is that cohorts differ in the environmental conditions experienced by the individuals during marking, such that subsequent survival, or recapture, or both, are affected.

In essence then, a cohort model is simply one where survival and/or recapture probabilities differ as a function of the cohort an animal is first released into. In its simplest form, a cohort model for survival (but **not** recapture) can be represented as:



For recapture, we still maintain the usual time-dependent parameter structure:



In this case, survival is constant over time, but differs among cohorts. How would this be represented by **MARK**? In other words, what would the PIMs look like?

For survival,

1	1	1	1
	2	2	2
		3	3
			4

and for recapture

5	6	7	8
	6	7	8
		7	8
			8

Of course, we could add time-dependence within-cohort, but this will substantially increase the number of parameters. For example, consider the following (cohort \times time) model for survival (shown at the top of the next page) – estimates differ over time *within* cohort, but over a given interval, they differ *among* cohorts.

cohort 1	1	$\xrightarrow{\varphi_1}$	2	$\xrightarrow{\varphi_2}$	3	$\xrightarrow{\varphi_3}$	4	$\xrightarrow{\varphi_4}$	5
cohort 2			2	$\xrightarrow{\varphi_5}$	3	$\xrightarrow{\varphi_6}$	4	$\xrightarrow{\varphi_7}$	5
cohort 3					3	$\xrightarrow{\varphi_8}$	4	$\xrightarrow{\varphi_9}$	5
cohort 4							4	$\xrightarrow{\varphi_{10}}$	5

Not only is the number of parameters increased significantly, but the number of estimable and non-estimable parameters is also changed. For example, suppose that the preceding model is applied to survival, but simple time-dependence for recaptures.

cohort 1	1	\rightarrow	2 p_2	\rightarrow	3 p_3	\rightarrow	4 p_4	\rightarrow	5 p_5
cohort 2			2	\rightarrow	3 p_3	\rightarrow	4 p_4	\rightarrow	5 p_5
cohort 3					3	\rightarrow	4 p_4	\rightarrow	5 p_5
cohort 4							4 p_4	\rightarrow	5 p_5

How many estimable parameters would there be? As discussed in Chapter 4, one way to count the number of estimable parameters for any single-group model is to write out the probability statements for each of the ‘saturated’ capture histories.

For this example, the saturated histories, and their corresponding probability statements are:

<i>encounter history</i>	<i>probability</i>
11111	$\varphi_1 p_2 \varphi_2 p_3 \varphi_3 p_4 \varphi_4 p_5$
01111	$\varphi_5 p_3 \varphi_6 p_4 \varphi_7 p_5$
00111	$\varphi_8 p_4 \varphi_9 p_5$
00011	$\varphi_{10} p_5$

How many unique, identifiable parameters are there? The key to answering this question lies in the terminal product terms (i.e., $\varphi_4 p_5, \varphi_7 p_5 \dots$). Each of these β terms contains p_5 . Is p_5 estimable? No – not without more information! As such, we have 4 different β terms – and thus 13 parameters.

As you might imagine, with progressively complex models, it can be more work to count the number of parameters. And, knowing the number of parameters is essential for model selection.

begin sidebar

cohort models with individuals marked as young

What about cohort models where individuals are marked as young? Well, in such cases, you run into an interesting consideration: for animals marked as young, as time passes, they get older. Thus, within a cohort, age and time are the same (as we've discussed previously). In fact, age, time and cohort are collinear, since

$$\text{age} = \text{time} - \text{cohort}$$

As such, you **can't** consider all 3 factors simultaneously (i.e., you can look at age and cohort, or age and time, or cohort and time, but not all 3 together).

And, in fact, this may be the larger reason why use of 'cohort' models is relatively infrequent. The most common scenario where you might be interested in a 'cohort' effect is when you have reason to believe (or evidence suggesting) that early development (growth) strongly influence future survival (say). Individuals brought up in 'good years' might have higher survival than individuals brought up in 'bad years'. This is more than likely to occur through environmental influences (i.e., 'good' vs. 'bad' years) on early growth (i.e., over the time of development from juvenile → adult).

the problem, then, is how do you disentangle 'cohort' effects from the influence of both 'age' and 'time', when the 3 factors are collinear? The answer is, you can't, as already noted earlier. You can evaluate at most 2 of 3 factors, which is particularly problematic in this type of scenario, since you're invariably leaving out something important.

However, for the sake of completeness, we forge on...

end sidebar

7.6.1. Building cohort models: PIMS and design matrices

The preceding suggests pretty strongly that in most cases, you'll build cohort models using PIMs. But, as you by now no doubt realize, if you want to apply any constraints to the model, or build additive models, then you'll need to use the design matrix.

Although the basic ideas behind building design matrices for cohort models are pretty much the same as we've seen before, you need to think a bit when it comes to (cohort × time) models. Let's have a quick look at a couple of them.

Suppose you have a 5 occasion mark-recapture study, and you are interested in building a series of cohort models for survival. Suppose your candidate model set consists of $\{\varphi_{cohort}\}$ and $\{\varphi_{cohort.time}\}$.

The basic PIM structure for φ_{cohort} (i.e., cohort effects only, no time variation) is

1	1	1	1
	2	2	2
		3	3
			4

whereas the PIM structure for $\varphi_{cohort.time}$ is

1	2	3	4
	5	6	7
		8	9
			10

Now, there are a couple of equivalent design matrices for both of these models. Recall from Chapter 6 that we can often use the identity matrix to code for certain linear models in place of the intercept-based approach. The advantage of the latter, however, is that it allows for easy testing of certain constrained models.

For the first model φ_{cohort} , the two equivalent design matrices would be

$$\mathbf{X}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{X}_2 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

The first design matrix (\mathbf{X}_1) is the standard identity matrix. In the second design matrix (\mathbf{X}_2), the first column is the intercept, while the next 3 columns code for the level of the ‘treatment’, cohort. Since there are 4 cohorts in a 5 occasion study, then we need 3 columns. Each row of the matrix corresponds to a different cohort.

OK, probably pretty straightforward, at this point. What about model $\{\varphi_{cohort.time}\}$? Well, as it turns out, this is somewhat more complicated, for some of the same reasons building the design matrix for the fully time-dependent age model was a few pages back: because there are some ‘cohort \times time’ interactions which do not occur. As such, it turns out there are several equivalent intercept-based linear models that will yield the same fit. We’ll spend just a moment here on this – if for no other reason that it forces us to think a bit more deeply about design matrices.

Now, for model $\varphi_{cohort.time}$, the (10×10) identity matrix design structure corresponding to the PIM would look like:

1	2	3	4
	5	6	7
		8	9
			10

$$\mathbf{D} = \begin{bmatrix} d_{1,1} & 0 & \dots & d_{1,10} \\ 0 & d_{2,2} & \dots & d_{2,10} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_{10,10} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

What about the intercept-based design matrix? Obviously, we need a column for the intercept. The cohort and time columns should be pretty straightforward: we have 4 cohorts, so 3 columns: let '1 0 0' be the coding for cohort 1, '0 1 0' be the coding for cohort 2, '0 0 1' be the coding for cohort 3, and '0 0 0' be the coding for cohort 4.

But, recall that the number of time intervals (parameters) for each cohort decrements by one with each successive cohort (i.e., for cohort 1 there are 4 intervals, for cohort 2 there are 3 intervals, for cohort 2 there are 2 intervals, and for cohort 4, there is only the final interval).

So, thus far, our design matrix looks like

$$\begin{bmatrix} 1 & 1 & 0 & 0 & \dots \\ 1 & 1 & 0 & 0 & \dots \\ 1 & 1 & 0 & 0 & \dots \\ 1 & 1 & 0 & 0 & \dots \\ 1 & 0 & 1 & 0 & \dots \\ 1 & 0 & 1 & 0 & \dots \\ 1 & 0 & 1 & 0 & \dots \\ 1 & 0 & 0 & 1 & \dots \\ 1 & 0 & 0 & 1 & \dots \\ 1 & 0 & 0 & 0 & \dots \end{bmatrix}$$

Now, for time, the approach is similar: we have 5 occasions (4 intervals), so therefore, we need 3 columns. We let '1 0 0' be the first time interval, '0 1 0' be the second time interval, '0 0 1' be the third interval, and '0 0 0' the final time interval. This is the usual 'reference cell' coding approach we discussed at length in Chapter 6.

Adding these time columns to the design matrix gives us

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & \dots \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & \dots \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & \dots \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & \dots \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & \dots \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & \dots \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \end{bmatrix}$$

How big does the design matrix need to be? Well, there are 10 parameters, so the full $\varphi_{\text{cohort} \times \text{time}}$ model corresponds to a (10×10) design matrix. Thus, we have 3 more columns to complete. OK, now the challenge. How can we code an interaction of (cohort \times time) in only 3 columns, if there are 4 cohorts!

Normally, we think of the interaction term(s) as products of the main factors in the linear model: in this case, the product of the 3 columns for 'cohort', and the 3 columns for 'time', which would yield 9 columns, not 3! But, if you look closely at the PIM (below), you'll see that not all interactions are possible. This is exactly the same issue we faced earlier when we were looking at interactions in models for analysis of data of individuals marked as both young, and adults – not all interactions were possible.

Here again is the PIM for model $\{\varphi_{\text{cohort} \times \text{time}}\}$:

1	2	3	4
	5	6	7
		8	9
			10

We see that for cohort 1, time interval 1 does not interact with any other cohort. We also see that cohorts 1 and 2 interact in both time interval 2, interval 3, and time interval 4. We see that cohort 3 only interacts with any other cohort (cohorts 1 and 2) in time intervals 3 and 4. For cohort 4, there appears to be an interaction in time interval 4, but, parameter 10 is not identifiable, so in fact, it doesn't really interact with anything!

So, we really have only 2 interaction blocks (indicated in the shaded 'blocks' in the following):

1	2	3	4
	5	6	7
		8	9
			10

1	2	3	4
	5	6	7
		8	9
			10

As it turns out, you only need to code either of these 'interaction blocks' to get the appropriate model fit (and parameter estimates).

So, for example, for the second of these two 'interaction blocks':

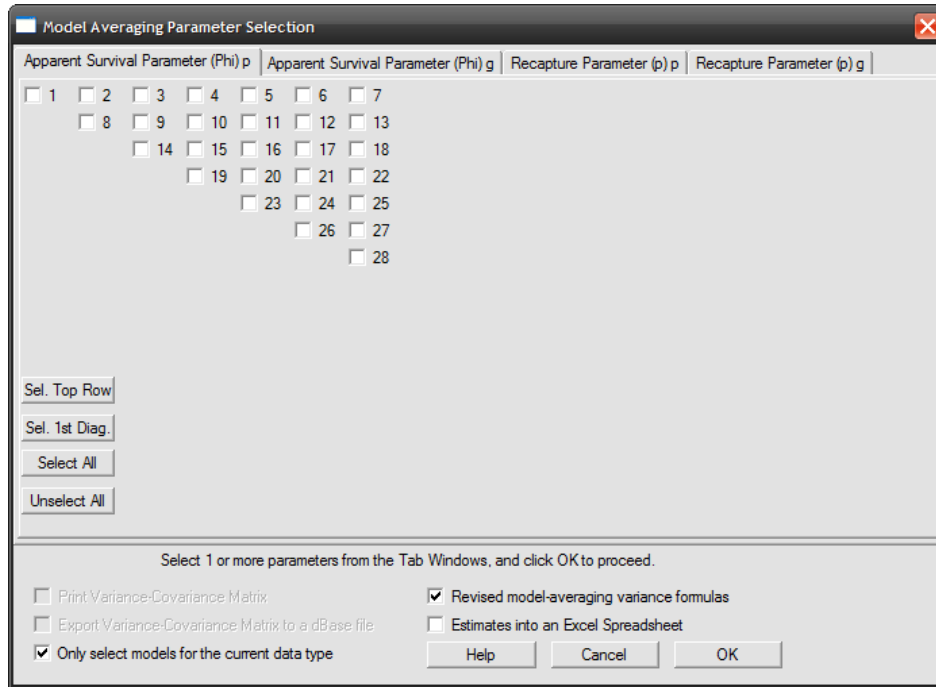
$$\begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Try it and see! A bit tricky at first, but hopefully you'll see the basic logic. Again, full (cohort \times time) models are relatively rare, owing to significant identifiability problems, but constrained models, especially where constraints are applied within cohort, are not uncommon. And to do this, you need to be able to build the design matrix for the full (cohort \times time) model(s) first.

7.7. Model averaging and age/cohort models

In chapter 4 we introduced the important concept of 'model averaging'. However, as you may recall, there were a few parts of the description of the 'mechanics' of model averaging in **MARK** that were somewhat 'obtuse'. That's because back in chapter 4 we had to make vague references to these things called 'age models' and 'cohort' models. New stuff in chapter 4, but by this point you're expert in both concepts, so we can re-visit some of the 'model averaging mechanics' and clear up a few points.

First, recall that for the swift data set, the model averaging window looked like the following:



For that data set, there were 2 groups (good colony, poor colony), and two parameters (survival probability, and recapture probability). Thus, 4 tabs along the top of the model averaging window (one for each group and parameter combination). In addition, for the swift data set, there were 8 occasions, so 7 intervals for survival.

Now, if we look carefully at the ‘triangular matrix’ structure that the model averaging window presents us with (above), we see something that at this point should be a bit more understandable. What would a full (time \times cohort) PIM look like for a data set with 8 occasions? Answer: just like the triangular matrix pictured above!

To make sure you understand model averaging, let’s analyze some simulated data (contained in `age2_avg.inp`). These simulated data consist of 7 occasions, 2 age-classes for survival, and a single age-class for encounter probability. We’ll fit 4 approximating models to these data: $\{\varphi_{t/t} p_t\}$, $\{\varphi_{t+t} p_t\}$, $\{\varphi_{t/.} p_t\}$, and $\{\varphi_{./} p_t\}$. Note that model $\{\varphi_{t+t} p_t\}$ has additive temporal effects between the 2 age classes, so we’ll construct all 4 models using a design matrix approach.

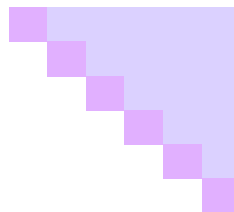
The results from our model fitting are shown at the top of the next page. We see that the most parsimonious model (model $\{\varphi_{t/.} p_t\}$) – time-varying survival for the first age class, time-constant survival for the second age-class, is a bit less than twice as well supported as the next best model ($0.636/0.346 = 1.84$).

Model	AICc	Delta AICc	AICc Weight	Model Likelihood	No. Par.	Deviance	-2Log(L)
{phi(a2 t.)p(t) - DM}	1577.3476	0.0000	0.63617	1.0000	12	90.7811	1552.9638
{phi(a2 t+t)p(t) - DM}	1578.5666	1.2190	0.34584	0.5436	12	92.0002	1554.1828
{phi(a2 tt)p(t) - DM}	1584.4905	7.1429	0.01789	0.0281	16	89.6354	1551.8181
{phi(a2 .)p(t) - DM}	1594.7531	17.4055	0.00011	0.0002	8	116.3941	1578.5768

What are the model averaged survival values for each interval for the first age-class? Again, remember that the first age-class is one-year in duration. To derive the model averaged survival probabilities for the first age-class, select '**Model Averaging**' from the '**Output**' menu in **MARK**. Note that this time, the model averaging window (shown below) has only 2 tabs: for survival and recapture, respectively. The triangular matrix (which you now recognize as the corresponding (time \times cohort) PIM structure) that the model averaging window presents is indexed from 1 to 21. The only thing you need to do at this stage is check the elements of this matrix corresponding to the first-year survival probabilities.

What would they be? Yes! – the elements along the diagonal – {1, 7, 12, 16, 19, 21} – as shown below:

These correspond to the diagonal elements of the following PIM structure:



Once you've satisfied yourself that this is correct – and makes sense – then click the '**OK**' button. The model averaged survival values are:

Parameter	Estimate	SE	LCI	UCI
Apparent Survival Parameter (Phi) Group 1 Parameter 1	0.3052793	0.0570497	0.20595	0.426768
Apparent Survival Parameter (Phi) Group 1 Parameter 7	0.5959166	0.065298	0.464307	0.715037
Apparent Survival Parameter (Phi) Group 1 Parameter 12	0.5039241	0.0695703	0.370571	0.636721
Apparent Survival Parameter (Phi) Group 1 Parameter 16	0.4121025	0.0720444	0.28128	0.556646
Apparent Survival Parameter (Phi) Group 1 Parameter 19	0.463646	0.0888382	0.300301	0.635184
Apparent Survival Parameter (Phi) Group 1 Parameter 21	0.9916067	8.9888927	-16.6266	18.60984

What about 'adult' survival? Again, the only trick is to remember that adult survival is 'above the diagonal' (look at the PIM schematic on the preceding page; the adult 'elements' are the shaded elements above the diagonal). So, you simply need to check the appropriate elements of the triangular PIM structure in the model averaging window.

But do you need to check all of them (i.e., 2 to 6, 8 to 11, 13 to 15, 17 to 18 and 20)? If we remind you that our models have time-structure only, but no cohort structure, you should realize with a bit of

thought that the answer is ‘no’. You simply need to check at least one of the ‘above diagonal’ elements in each column (each column where there is an above-diagonal element).

Thus, the following two model averaging windows shown below

Apparent Survival Parameter (Phi) Group 1					
<input type="checkbox"/> 1	<input checked="" type="checkbox"/> 2	<input checked="" type="checkbox"/> 3	<input checked="" type="checkbox"/> 4	<input checked="" type="checkbox"/> 5	<input checked="" type="checkbox"/> 6
	<input type="checkbox"/> 7	<input type="checkbox"/> 8	<input type="checkbox"/> 9	<input type="checkbox"/> 10	<input type="checkbox"/> 11
		<input type="checkbox"/> 12	<input type="checkbox"/> 13	<input type="checkbox"/> 14	<input type="checkbox"/> 15
			<input type="checkbox"/> 16	<input type="checkbox"/> 17	<input type="checkbox"/> 18
				<input type="checkbox"/> 19	<input type="checkbox"/> 20
					<input type="checkbox"/> 21

OR

Apparent Survival Parameter (Phi) Group 1					
<input type="checkbox"/> 1	<input checked="" type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6
	<input type="checkbox"/> 7	<input checked="" type="checkbox"/> 8	<input type="checkbox"/> 9	<input checked="" type="checkbox"/> 10	<input type="checkbox"/> 11
		<input type="checkbox"/> 12	<input checked="" type="checkbox"/> 13	<input type="checkbox"/> 14	<input checked="" type="checkbox"/> 15
			<input type="checkbox"/> 16	<input type="checkbox"/> 17	<input type="checkbox"/> 18
				<input type="checkbox"/> 19	<input type="checkbox"/> 20
					<input type="checkbox"/> 21

will yield entirely equivalent results. Try it and see! Make sure you understand *why* they are equivalent.

Model averaging is an important technique, so make sure you really understand what’s going on here. The model averaging window also forces you to come to grips with your understanding of age, cohort, and time models. Remember, ‘age’ = (‘cohort’ – ‘time’). So, make sure you do ‘get it’.

7.8. Summary

That is the end of Chapter 7. We have considerably expanded the range of ‘underlying’ models we can fit to mark-recapture data, by developing age/TSM and cohort models. We have also seen how constraints can be applied to these models as easily as we did with CJS models. We also revisited the idea of model averaging. These models are **very** widely used (the basic structure shows up various ways for a number of different data types), so make sure you thoroughly understand the material in this chapter. In the next chapter we look at a very different data type – ‘dead recoveries’.

7.9. References

Pradel, R., Hines, J. E., Lebreton, J.-D., and Nichols, J. D. (1997) Capture-recapture survival models taking account of transients. *Biometrics*, **53**, 60-72.