

# 'Agegap' Formation hazard

October 10, 2016

The hazard used for the formation and dissolution events (between persons  $i$  and  $j$ ) has the following form:

$$\begin{aligned}
 h &= \exp(\alpha_0 + \alpha_1 P_i + \alpha_2 P_j + \alpha_3 |P_i - P_j| \\
 &+ \alpha_4 \left( \frac{A_i(t) + A_j(t)}{2} \right) \\
 &+ \alpha_5 |A_i(t) - A_j(t) - D_{p,i} - \alpha_8 A_i(t)| \\
 &+ \alpha_9 |A_i(t) - A_j(t) - D_{p,j} - \alpha_{10} A_j(t)| \\
 &+ \beta t_{diff})
 \end{aligned}$$

The properties used are:

- $\alpha_k, \beta$  : Constants
- $P_i, P_j$  : Number of partners of these persons
- $A_i, A_j$  : Age
- $D_{p,i}, D_{p,j}$  : Preferred age difference for persons  $i$  and  $j$
- $t_{diff}$ : Time since relationship became possible between these two persons

Here, both the age and  $t_{diff}$  contain a time dependency as follows:

$$A_i(t) = t - t_{B,i}$$

$$t_{diff} = t - t_r$$

where

- $t_{B,i}$  : Time at which person  $i$  was born
- $t_r$  : Time at which relationship became possible between these two persons

This hazard can be rewritten as follows, where the time dependency is now explicitly shown:

$$h = \exp(B + \alpha_4 t + \beta t + \alpha_5 |C - \alpha_8 t| + \alpha_9 |D - \alpha_{10} t|)$$

In this expression,  $B, C$  and  $D$  are no longer time dependent:

$$B = \alpha_0 + \alpha_1 P_i + \alpha_2 P_j + \alpha_3 |P_i - P_j| - \alpha_4 \left( \frac{t_{B,i} + t_{B,j}}{2} \right) - \beta t_r$$

$$C = (\alpha_8 - 1)t_{B,i} + t_{B,j} - D_{p,i}$$

$$D = (\alpha_{10} + 1)t_{B,j} - t_{B,i} - D_{p,j}$$

Because of the absolute values, there are four different situations we need to take into account:

1.  $C > \alpha_8 t$  and  $D > \alpha_{10} t$

$$h = \exp(B + \alpha_5 C + \alpha_9 D + t(\alpha_4 + \beta - \alpha_5 \alpha_8 - \alpha_9 \alpha_{10}))$$

2.  $C > \alpha_8 t$  and  $D < \alpha_{10} t$

$$h = \exp(B + \alpha_5 C - \alpha_9 D + t(\alpha_4 + \beta - \alpha_5 \alpha_8 + \alpha_9 \alpha_{10}))$$

3.  $C < \alpha_8 t$  and  $D > \alpha_{10} t$

$$h = \exp(B - \alpha_5 C + \alpha_9 D + t(\alpha_4 + \beta + \alpha_5 \alpha_8 - \alpha_9 \alpha_{10}))$$

4.  $C < \alpha_8 t$  and  $D < \alpha_{10} t$

$$h = \exp(B - \alpha_5 C - \alpha_9 D + t(\alpha_4 + \beta + \alpha_5 \alpha_8 + \alpha_9 \alpha_{10}))$$

For each of the cases the hazard is of the form

$$\exp(E + tF)$$

where in each of the cases:

1.  $E = B + \alpha_5 C + \alpha_9 D$      $F = \alpha_4 + \beta - \alpha_5 \alpha_8 - \alpha_9 \alpha_{10}$
2.  $E = B + \alpha_5 C - \alpha_9 D$      $F = \alpha_4 + \beta - \alpha_5 \alpha_8 + \alpha_9 \alpha_{10}$
3.  $E = B - \alpha_5 C + \alpha_9 D$      $F = \alpha_4 + \beta + \alpha_5 \alpha_8 - \alpha_9 \alpha_{10}$
4.  $E = B - \alpha_5 C - \alpha_9 D$      $F = \alpha_4 + \beta + \alpha_5 \alpha_8 + \alpha_9 \alpha_{10}$

The relevant integral for a hazard of this form is

$$\Delta T = \int_{t_0}^{t_0 + \Delta t} \exp(E + Ft) dt$$

which has the analytical solution

$$\Delta T = \frac{\exp(E + Ft_0)}{F} (\exp(F \Delta t) - 1)$$

and inverse

$$\Delta t = \frac{1}{F} \ln \left[ \frac{F \Delta t}{\exp(E + Ft_0)} + 1 \right]$$

However, because of the absolute values in the integral, there are in general two times at which a sign inside an absolute value changes: one is when  $C = \alpha_8 t$ , the other when  $D = \alpha_{10} t$ . Calling the earliest of these times  $t_1^P$  and the other one  $t_2^P$ , the mapping from  $\Delta t$  to  $\Delta T$  could require three different forms of the hazard:

$$\Delta T = \int_{t_0}^{t_0 + \Delta t} h(t) dt = \int_{t_0}^{t_1^P} h_1(t) dt + \int_{t_1^P}^{t_2^P} h_2(t) dt + \int_{t_2^P}^{t_0 + \Delta t} h_3(t) dt$$

Depending on the precise value of  $\Delta t$ , it could also be simpler than this.

Calculating  $\Delta T$  from  $\Delta t$  is the most straightforward part, since the relative position of  $t_0 + \Delta t$  with respect to  $t_1^P$  and  $t_2^P$  is known. When starting from  $\Delta T$  and solving for  $\Delta t$  this is no longer the case.

To solve for  $\Delta t$ , we first need to calculate

$$\Delta T_1^P = \int_{t_0}^{t_1^P} h_1(t) dt$$

If this value is larger than  $\Delta T$ , we know that the value of  $t_0 + \Delta t$  must be smaller than  $t_1^P$  and we simply need to solve

$$\Delta T = \int_{t_0}^{t_0 + \Delta t} h_1(t) dt$$

If the value is smaller than  $\Delta T$ , we know that  $t_0 + \Delta t$  is larger than  $t_1^P$  and we need to calculate

$$\Delta T_2^P = \int_{t_1^P}^{t_2^P} h_2(t) dt$$

If  $\Delta T$  is smaller than  $\Delta T_1^P + \Delta T_2^P$ , we know that the value of  $t_0 + \Delta t$  is smaller than  $t_2^P$  and we need to solve

$$\Delta T - \Delta T_1^P = \int_{t_1^P}^{t_0 + \Delta t} h_2(t) dt$$

In the final case, if  $\Delta T$  is larger than  $\Delta T_1^P + \Delta T_2^P$ , we need to solve

$$\Delta T - \Delta T_1^P - \Delta T_2^P = \int_{t_2^P}^{t_0 + \Delta t} h_3(t) dt$$