

Some notes on R code and Overleaf Likelihood Expressions

10 January 2023

OL equation 7 - d_fit_sus_cens_posttest

January 10, 2023

This one is very straightforward and the likelihood expressions in Overleaf looks correct to me. And the R code looks correct, and it passes the test.

OL equation 9 - d_fit_sus_cens_postno and d_fit_endlive

January 10, 2023

Overleaf text

The description under 3.2.1 says "...were right censored and *were* tested at censoring." It should say "... *were not* tested..."

Equation 8 - Continuous time formulation

The index should be e_i , not r_i . That is, $\exp\left(-\int_{r_i}^{d_{n-1}} \lambda_S(y) dy\right)$ should be $\exp\left(-\int_{e_i}^{d_{n-1}} \lambda_S(y) dy\right)$.

Also, the FOI piece is missing: I think the second line, inside the square brackets, should have $\times \exp\left(-\int_{d_{n-1}}^x \lambda_F(y) dy\right)$.

Equation 9 - Unit cumulative hazard approximation

As in Eq 8, the FOI piece is missing. In addition, the susceptible survival sum index should begin at $d_{n-1} + 1$ because line 1 takes care of the d_{n-1} case. So, to fix both issues, in the second line $\exp\left(-\sum_{y=d_{n-1}}^{x-1} \Lambda_S(y)\right)$ should be replaced with

$$\exp\left(-\sum_{y=1}^{x-1} \Lambda_F(y)\right) \times \exp\left(-\sum_{y=d_{n-1}+1}^{x-1} \Lambda_S(y)\right)$$

R code

In the preliminary hazards, the infected survival piece has `lam_inf[1:(r - 1)]` and `age_effect_surv[1:(r - 1)]`. I think in both cases the indices should be `e:(r-1)`. Also, the line `period_effect_surv[(1 + age2date):(r - 1 + age2date)]` either should be `period_effect_surv[(e + age2date):(r - 1 +`

age2date)]. But actually, I guess this does not even matter, because the likelihood then uses only the values from e onward, and ignores everything else. But it seems unnecessary to populate values for $1:(e-1)$.

Down in the joint likelihood bit, I think the following

```
for(k in (e + 2):(r - 1)){
  lik_temp[k] <- lam_foi[k] *
    exp(-sum(lam_sus[e:(k - 1)])) *
    exp(-sum(lam_inf[k:(r - 1)]))
}
```

needs to be replaced with

```
for(k in (e + 2):(r - 1)){
  lik_temp[k] <- lam_foi[k] *
    exp(-sum(lam_sus[e:(k - 1)])) *
    exp(-sum(lam_foi[e:(k - 1)])) *
    exp(-sum(lam_inf[k:(r - 1)]))
}
```

Also, I know that in this case $d_{n-1} = e_i$, but in the R code for other likelihoods, `dn1` is used instead of `e`, e.g., `for(k in (dn2 + 2):(r - 1))` instead of `for(k in (e + 2):(r - 1))`.

The distribution test passes for the `d_fit_sus_cens_postno` dataset, but it fails in every case for the `d_fit_endlive` dataset. The reason it fails is that `r-1+age2date = 1563`, but the `period_effect_surv` part of the hazard functions has length 1562. The `r-1+age2date = 1563` bit does not cause a problem in the FOI hazard, because `period_lookup`, which is used to define `indx_foi_period`, has length 1604.

OL equation 13 - dNegCap_MortNoTest - UPDATE

January 10, 2023

As I was working on understanding Overleaf equation 9, I took another look at this one, and realized that there are still two mistakes (I think) in the Unit Cumulative Hazard Approximation. First, the one square bracket in line 2 is at the wrong place (It was in the right place until I convinced AK to move it!). Second, also in the second line, I think the FOI summation index should go to $x - 1$ instead of x , because the individuals cannot escape infection to x , and then also get infected at x . So I think that in the second line, the following

$$\exp\left(-\sum_{y=d_{n-1}+1}^x \Lambda_F(y)\right) \times \left[\exp\left(-\sum_{y=d_{n-1}+1}^{x-1} \Lambda_S(y)\right)\right]^{I[x>d_{n-1}+1]}$$

should be replaced with

$$\left[\exp\left(-\sum_{y=d_{n-1}+1}^{x-1} \Lambda_F(y)\right) \times \exp\left(-\sum_{y=d_{n-1}+1}^{x-1} \Lambda_S(y)\right)\right]^{I[x>d_{n-1}+1]}$$

OL equation 21 - dNegCap_NegRecap_Cens_PostNo - UPDATE

January 10, 2023

As I was working on understanding Overleaf equation 9, I took another look at this one, I think I convinced myself that in both the integral and the summation, there is a piece missing for the FOI for those that get infected between d_{n-1} and r_i . In Equation 8, I think we need a $\exp\left(-\int_{d_{n-1}}^x \lambda_F(y) dy\right)$ inside the square brackets.

And in the unit cumulative hazard approximation, there are two problems. We also need the FOI piece in line 2, and also in line 2, I think the susceptible survival bit should be indexed from $d_{n-1} + 1$. SO, if I am right, the following expression

$$\left(\exp \left(- \sum_{y=d_{n-1}}^{x-1} \Lambda_S(y) \right) \right)^{I[x > d_{n-1} + 1]}$$

should be replaced with

$$\left(\exp \left(- \sum_{y=d_{n-1}+1}^{x-1} \Lambda_F(y) \right) \times \exp \left(- \sum_{y=d_{n-1}+1}^{x-1} \Lambda_S(y) \right) \right)^{I[x > d_{n-1} + 1]}$$

R code

think the FOI bit needs to be added into the R code.

```
for(k in (dn1 + 2):(r - 1)){
  lik_temp[k] <- lam_foi[k] *
    exp(-sum(lam_sus[dn1:(k - 1)])) *
    exp(-sum(lam_foi[dn1:(k - 1)])) *   ### add this line
    exp(-sum(lam_inf[k:(r - 1)]))
}
```

OL equation 29 - dNegCapPosMort - UPDATE

January 10, 2023

As I was working on understanding Overleaf equation 9, I also took another look at Eq. 13 and this one. This one should be identical to 13, except it removes the segment that allows individuals to die uninfected. But 13 contains a piece for the case of d_{n-1} that is currently missing here, and I think needs to be included. Also, I think there are still 2 additional issues with the approximation (same as for 13). First, I think the summation index for FOI in line 2 should go only to $x - 1$ rather than x . Also, I think that FOI summation should therefore go inside the square brackets with the Λ_S bit.

Equation 28

For Equation 28, I think the first line should be

$$\exp \left(- \int_0^{d_{n-1}} \lambda_F(y) dy \right) \times \exp \left(- \int_{e_i}^{d_{n-1}} \lambda_S(y) dy \right)$$

rather than

$$\left(1 - \exp \left(- \int_{r_i}^{s_i} \lambda_I(y) dy \right) \right)$$

Equation 29

Similarly, for Equation 29, I think the first line (like in Equation 13) should be

$$\exp\left(-\sum_{y=1}^{d_{n-1}} \Lambda_F(y)\right) \times \exp\left(-\sum_{y=e_i}^{d_{n-1}} \Lambda_S(y)\right)$$

rather than

$$\left(1 - \exp\left(-\sum_{y=r_i}^{s_i-1} \Lambda_I(y)\right)\right)$$

Also, just like with 13, I think the following code in line 2

$$\exp\left(-\sum_{y=d_{n-1}+1}^x \Lambda_F(y)\right) \times \left[\exp\left(-\sum_{y=d_{n-1}+1}^{x-1} \Lambda_S(y)\right)\right]^{I[x>d_{n-1}+1]}$$

should be replaced with

$$\left[\exp\left(-\sum_{y=d_{n-1}+1}^{x-1} \Lambda_F(y)\right) \times \exp\left(-\sum_{y=d_{n-1}+1}^{x-1} \Lambda_S(y)\right)\right]^{I[x>d_{n-1}+1]}$$

R code

Again, this should be analogous to the code for 13, so a few changes are needed.

```
for (k in (r - 1):(s - 2)) {
  lik_temp[k] <- lam_foi[k] *
    exp(-sum(lam_foi[(dn1 + 1):k])) *
    exp(-sum(lam_sus[(dn1 + 1):(k - 1)])) *
    exp(-sum(lam_inf[k:(r - 1)])) *          ### remove this line
    (1 - exp(-sum(lam_inf[k:(s - 1)])))
}
lik_temp[(s - 1)] <- lam_foi[(s - 1)] *
  exp(-sum(lam_foi[(dn1 + 1):(s - 1)])) *
  exp(-sum(lam_sus[(dn1 + 1):((s - 1) - 1)])) *
  ((1 - exp(-lam_inf[(s - 1)])) + lam_inf[(s - 1)]) ### replace this line
  (lam_inf[(s - 1)])                                ### with this line

### replace this
lik <- (1 - exp(-sum(lam_sus[r:(s - 1)]))) *
  sum(lik_temp[(dn1 + 1):(s - 1)])
### with this
lik <- exp(-sum(lam_sus[e:dn1])) *
  exp(-sum(lam_foi[1:dn1])) *
  sum(lik_temp[(dn1 + 1):(s - 1)])
```