Blantyre ACF Appendix 2

Pete Dodd

2021-11-16

Mathematical formulation of models

Let $c \in \{1, 2\}$ index the population, with 1 corresponding to the ACF population and 2 to the 'control' population. I think it helps in setting up counterfactuals to keep the population index separate from any indices that control whether ACF is applied.

Let t measure time in quarters from some natural reference point.

Let $\alpha \in \{ACF, notACF\}$ denote whether ACF is applied or not.

Let $P_{c,t}$ be the populations at each time, and let $\pi_{c,t} = \log(P_{c,t})$.

Let $\ell_{c,t,\alpha}$ be the corresponding Poisson rate parameter on a log scale. Let $r_{c,t,\alpha} = \exp(\ell_{c,t,\alpha})$ be the Poisson rate itself.

We will write $\mathbb{I}(t)$ for an indicator function that is 1 during the ACF period, and 0 otherwise, and δ_α for an indicator function that is 1 when $\alpha = ACF$ and 0 otherwise.

Without control

Here, we only have data from the ACF population, and some notation is redundant since this means $c = 1$.

$\ell_{c,t,\alpha} = \pi_{c,t} + k_c + s_c.t + \delta_\alpha \mathbb{I}(t)(a + b.t)$

Here k is the intercept and s the slope, and a and b represent the respective increments to these under ACF. In $\alpha = ACF$ corresponds to the process that gave rise to the data from population 1.

With control

We now want to capture both the intervention effect during the ACF period, and a non-intervention effect during the ACF period. Population 2 allows estimation of the latter. We can separate these out explicitly:

$\ell_{1,t,\alpha} = \pi_{1,t} + k_1 + s.t + \mathbb{I}(t)(\delta_\alpha[a + b.t] + [A + B.t])$

$\ell_{2,t,\alpha} = \pi_{2,t} + k_2 + s.t + \mathbb{I}(t)(A + B.t)$

In fitting to data, $\alpha = ACF$ for $c = 1$ and $\alpha = notACF$ for $c = 2$.

Note: we have restricted the slope (s) in each population prior to the ACF period to be equal due to model fits lacking face validity.

Definition of quantity of interest

Having fitted the models, we want to compute for each the expected cumulative difference in notifications between no-ACF and ACF conditions for the ACF community, $D(\theta)$. This is a function of the model parameters which we will collectively denote θ. That is

$$D(\theta) = \sum_t \mathbb{I}(t)[r_{1,t,ACF} - r_{1,t,noACF}]$$

If $t = t_1$ is the first time in the ACF period, and $t = t_2$ the last,

$$D(\theta) = e^{k_1+a} \sum_{t=t_1}^{t_2} P_{1,t} \left(e^{(s_1+b).t} - e^{b.t} \right)$$

which has no closed-form answer with the population offset.

For the no-control approach, the corresponding formula is

$$D_{wc}(\theta) = \sum_{t=t_1}^{t_2} P_{1,t} e^{k_1+s.t} \left(e^{\bar{a}+\bar{b}.t} - e^{A+B.t} \right)$$

where $\bar{a} = a + A$ and $\bar{b} = b + B$.

If the estimate for θ is asymptotically normal with mean $\bar{\theta}$ and variance-covariance matrix Σ, an approximation is that $D(\theta)$ is asymptotically normal with mean $D(\bar{\theta})$ and variance-covariance $J^T \Sigma J$, where J is the gradient (derivative) of D with respect to the parameters θ.

Availability of code

The supplementary data and code files (available at www.github.com/rachaelmburke/XXXX) show the implementation of this work in R.