Contents

1. Basic Model

1.1 Basic Terms

- Attack rate = $\frac{\text{affected people}}{\text{overall population}}$
- serial interval = some way to model time needed from infection to disease outbreak
- incidence = number of new cases reported
- Reproduction number R_0 = average number of people infected by 1 person. The reproduction number is also the expected value of the
- Offspring distribution = the distribution of the number of new infections from each infected person.
- force of infection = expected number of incidences at a certain time, i.e. how many infected people do you expect to be infected the next day.

1.2 Model from the other master thesis

the force of infection is modeled as

$$E(I_t) = \lambda_t = R_t \sum_{s=1}^{t-1} I_s \cdot \omega_{t-s} \tag{1.1}$$

with

- I_t = incidence at time t
- ω_{t-s} weighting parameter that models the serial interval, i.e. the development of the disease in one person
- R_t = the current Reproduction number, i.e. the number of people that are infected by one person.

In summary, the current incidence is modeled as the total number of affected people in period t-1 times the Reproduction rate in period t. The number of affected people in period t-1 is computed as the sum of all past incidences multiplied by a weight that models the course / progression of the disease (i.e. if people are dead after 10 days with 50% probability, then the 10-day lag should be weighted with 0.5).
One can then model each of the model parts with its own distribution:

- the *serial interval* is modeled as a gamma distribution gamma(2.706556, 0.1768991) with mean 15.3 and sd 9.3. days. This must then be discretized because we only model daily time steps.
- the *offspring distribution* can most easily be modeled as the Poisson distribution $P(X = k) = e^{-\lambda}\frac{\lambda^k}{\lambda!}$ or as the negative binomial distribution with mean μ and variance $\mu + \frac{\mu^2}{k}$ with dispersion parameter k. A smaller k means that the disease outbreak is dominated by a few super-spreaders, while a larger k means that all people infect others similarly.

1.3 Estimation of the Model

1.3.1 First Option

To predict the incidence I_t one first needs to predict the reproduction number R_t. Basically they compute all parameters to get a posterior distribution of R_{t-1} and then draw samples from this posterior distribution to obtain an estimated value R_t (assuming the value stays on average constant).

1.3.2 Second Option

Relaxing the assumption that the value R_t stays constant from R_{t-1}, they apply a Bayesian structural time series model.

2. Assessing Probabilistic Models

2.1 Calibration

Calibration means that the forecasted distribution is equal to the actual distribution of observed values. If you predict it will rain with 60% probability you should see rain in 60% of cases.

2.1.1 Assessing Calibration

One can use a probability integral transformation (PIT). You compare a predicted CDF F with a true CDF G by computing

$$u_t = F_t(k_t) - \nu(F_t(k_t) - F_t(k_t - 1)) \tag{2.1}$$

where ν is a uniform random variable between 0 and 1. If the prediction is ideal, the values u_t will be standard uniformly distributed. $\Rightarrow$ test uniform distribution with Anderson-Darling test. Test over- or underdispersion of estimated distribution (i.e. whether your forecast over- or underestimates variation) by looking at the histogram of PIT values. If values cluster at the centre, then the forecast is overdispersed and variance is overestimated (and vice versa for values clustering at the edges). Formal measure of centrality:

$$centrality = \frac{\text{Number of } u_t \text{ values between 0.25 and 0.75}}{\text{Number of all values}} - 0.5$$

Score < 0 means the forecast underestimates the true variability. Score above 0 overestimates it.

2.2 Sharpness

Sharpness is defined as the range of values in the forecast and measures how certain a forecast is. The Sharpness therefore does not depend on the true values, but solely considers the forecast.

2.2.1 Assessing Sharpness

Sharpness can e.g. be assessed by looking at the normalised absolute deviation about the median of I_t.

$$S(I_t) = \frac{1}{0.675}\text{median}(|1 - \text{median}(I_t)|) \tag{2.2}$$

The normalisation factor makes sure that $S(I_t)$ corresponds to the standard deviation if F_t is normal. ??

2.3 Bias

Bias can be assessed with

$$B_t(F_t, k_t) = 1 - (F_t(k_) - F_t(k_t - 1)) \tag{2.3}$$

2.4 Proper Scoring Rules

Proper scoring rules are used to rank probabilistic forecasts. The ideal forecast minimizes the proper scoring rule. Proper scoring rules take into account calibration and sharpness.

3. Encorporating Spatial Aspects

- gravity model? model force of infection from outside the health zone

3.1 Current Models

3.1.1 BSTS with a local linear trend

The time series is modeled as

$$\mu_t + 1 = \mu + t + \delta_t + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma_\mu),$$

$$\delta_t = \delta_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\delta).$$

The local linear trend model therefore assumes that both the mean of the time series as well as the slope follow random walks.

3.1.2 BSTS with a local linear trend and Student's-t-distributed errors

The model specification is very similar to the local linear trend model. The only change is that random errors now follow a t-distribution, which implies thicker tails and therefore more room for extreme changes.

$$\mu_t + 1 = \mu + t + \delta_t + \epsilon_t, \quad \epsilon_t \sim \mathcal{T}_{\nu_\mu}(0, \sigma_\mu),$$

$$\delta_{t+1} = \delta_t + \eta_t, \quad \eta_t \sim \mathcal{T}_{\nu_\delta}(0, \sigma_\delta).$$

ν_μ and ν_δ are parameters that determine the thickness of the tails.

3.1.3 BSTS with a semi-local linear trend

The semi-local linear trend model assumes that the mean of the time series moves according to a random walk with slope. The slope component is an AR1 process centred on a constant trend D. For any non-zero value ϕ, the slope will eventually become a constant number for long time horizons, resulting in a linear trend for the overall time series. Values of ϕ closer to one imply that the time series (or forecasts made by the model, respectively) will converge quicker to a purely linear trend. The time series is modeled as

$$\mu_t + 1 = \mu + t + \delta_t + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma_\mu),$$

$$\delta_{t+1} = D + \phi(\delta_t - D), \quad \eta_t \sim \mathcal{N}(0, \sigma_\delta).$$

3.1.4 BSTS with an AR1 and AR2 state component

The mean of the time series is modeled as an AR(1)-process (or AR(2), respectively). The AR(1) model looks like this:

$$\alpha_t = \phi_1 \alpha_{t-1} + \epsilon_{t-1}, \quad \epsilon_t \sim \mathcal{N}(0, \sigma_\delta).$$

The AR(2) model looks like this:

$$\alpha_t = \phi_1 \alpha_{t-1} + \phi_2 \alpha_{t-2} + \epsilon_{t-1}, \quad \epsilon_t \sim \mathcal{N}(0, \sigma_\delta).$$

***

3.2 Important Software and Packages

- EpiEstim
- BSTS (Bayesian structural time series models)

General

papers to read

1.