Assignment 2

Brian Sohn

1. Estimating Parameters
   a. Organic

$$O_t = O_0 + \mu t + \varepsilon_{o,t}$$

We will model both O_0 and μ as a positive real number using a non-informative prior. The error terms $\varepsilon_{o,t}$ will follow normal distributions with mean 0, but each of them will have a different variance. We will use hierarchical modeling to introduce correlation to these variances and set non-informative priors accordingly.

$$O_0 \sim \text{HN}(100)$$

$$\mu \sim \text{HN}(10)$$

$$\varepsilon_{0,t} \sim \text{N}(0, \sigma_{o,t}^2)$$

$$\sigma_{o,t} \sim \text{HN}(\nu_o)$$

$$\nu_o \sim \text{HN}(10)$$

Finally, our likelihood for organic registrations will be

$$O_t \sim \text{N}(O_0 + \mu t,\ \sigma_{o,t}^2)$$

The corresponding observations to this will be the 'organic' column in the dataset.

To avoid divergences in pymc, we will reparameterize sigmas into taus when coding this up.

   b. Paid

$$R_{i,t} = S_i \sum_{j=0}^{5} d_i^j \left(1 - exp\left(-\lambda_i A_{i,t-j}\right)\right) + \varepsilon_{i,t}$$

      i. Saturation

To ensure S_i is positive, we will model it as an exponent of some real number S_{i_0}. S_{i_0} will follow a normal distribution with mean $S_{channel\_k}$ and variance σ_S^2. Here, channel_k is the channel that each subchannel i belongs to. As a result, there will be 3 values of $S_{channel\_k}$ resulting in and 8 values of S_{i_0}. The variance of S_{i_0} is modelled to be equal for all subchannels.

$$S_i = \exp(S_{i\_0})$$

$$S_{i\_0} \sim \text{N}(S_{channel\_k}, \sigma_S^2)$$

$$S_{channel\_k} \sim \text{N}(0,1)$$

$$\sigma_S \sim \text{HN}(0,1)$$

ii. Tail Decay

The same idea as saturation is applied here, where we model d_i as a sigmoid of a real number d_{i_0} so that d_i lies between 0 and 1. We also use the same hierarchical modelling idea as saturation, by assuming each d_{i_0} follows a normal distribution with mean $d_{channel\_k}$, which will be equal for all subchannels within a channel. Variance of d_{i_0} is also modelled just as saturation.

$$d_i = 1 / (1+\exp(-d_{i\_0}))$$

$$d_{i\_0} \sim \mathrm{N}(d_{channel\_k}, \sigma_d^2)$$

$$d_{channel\_k} \sim \mathrm{N}(0,1)$$

$$\sigma_d \sim \mathrm{HN}(0,1)$$

iii. Diminishing Returns

Diminishing returns are modelled the exact same way tail decay is modelled.

$$\lambda_i = 1 / (1+\exp(-\lambda_{i\_0}))$$

$$\lambda_{i\_0} \sim \mathrm{N}(\lambda_{channel\_k}, \lambda_d^2)$$

$$\lambda_{channel\_k} \sim \mathrm{N}(0,1)$$

$$\sigma_\lambda \sim \mathrm{HN}(0,1)$$

iv. Error terms

Error terms are modelled similarly to error terms for organic registrations. Each $\varepsilon_{i,t}$ is distributed normally with mean 0 and heteroskedastic variance $\sigma_{i,t}^2$. They are hierarchically modelled using a higher level parameter ν_p.

$$\varepsilon_{i,t} \sim \mathrm{N}(0, \sigma_{i,t}^2)$$

$$\sigma_{i,t} \sim \mathrm{HN}(\nu_p)$$

$$\nu_p \sim \mathrm{HN}(10)$$

v. Likelihood

Each $R_{i,t}$ is modelled as

$$R_{i,t} \sim \mathrm{N}(S_i \sum_{j=0}^{5} d_i^j \left(1 - exp\left(-\lambda_i A_{i,t-j}\right)\right), \sigma_{i,t}^2)$$

The corresponding observation to these variables are the survey-based estimations of each subchannel's attribution.

We reparameterize sigma into tau when needed to avoid divergences in pymc.

c. Results

Using posterior means, we get the following estimates for each parameter.

O_0: 91.7

μ: 0.297

	tv	hulu	yt	fb	ig	snap	google	bing
S	10.534	12.198	9.128	12.203	13.647	12.522	9.236	8.559
lambda	0.352	0.261	0.370	0.385	0.383	0.374	0.515	0.517
d	0.234	0.222	0.225	0.122	0.085	0.115	0.150	0.163

2. Finding an Optimal Advertising Strategy

For each subchannel, an additional dollar (or minimal unit of money) spent in advertising at time t will bring us

$$\frac{\partial\ R_{i,t}+R_{i,t+1}+R_{i,t+2}+R_{i,t+3}+R_{i,t+4}+R_{i,t+5}}{\partial A_{i,t}}$$

new registrations. Let's call the expectation of this value as the marginal registration. This value can be computed since we know all the parameters constituting it. Now we can optimize advertising spending by the following process:

step 1) Start by setting $A_{i,t}$ = 0 for all i.

step 2) Compute marginal registration for all subchannels.

step 3) Allocate the next dollar to the subchannel with highest marginal registration. Let's call this subchannel j.

step 4) Compute marginal registration for j again. Then, go back to step 3 if you have advertising budget left for time t. Otherwise (i.e. if you run out of advertising budget), stop.

Note that this greedy algorithm will work because of diminishing returns.