Haensel AMS GmbH task

We had total of 104 records having 7 different spending channels we did our analysis for each channel across weeks along with revenues and trend.

Please note: Sine MCMC is a probabilistic model you may find different values for decay factor and ROIs due to that but Bayesian MMM results remain same. When you run the code values might defer slightly for channel beta and ROI but overall result remains same that is best performing channels always perform well and non-performing channels remain non-performing if you one to get same result every time you can you use random seed I've added that in my code as a comment you can uncomment it if you wish to get the same result each time.

1: How do you model spend carry over?

The model spends carry over using the adstock transformation. Adstocking is a approach to account for the delayed effect of advertising spend. It models the carryover effect by applying a decay factor to past advertising spends, ensuring that the impact of an ad is spread out over time rather than being concentrated in the period it was spent.

Here we didn't use same decay factor for each channel instead calculated the optimized decay factor for each channel using a minimization process to best fit the observed data. The optimal decay factors are found by minimizing the difference between the predicted and actual log revenue.

2: Explain your choice of prior inputs to the model?

Intercept:

Mean : 0,Standard Deviation (sigma): 10

Reason: A normal distribution with a mean of 0 and a large standard deviation reflects that we have no strong prior belief about the intercept's value. The large sigma ensures that the prior is not too informative and allows the data to play a significant role in determining the posterior.

Beta (Regression Coefficients):

Mean: 0 Standard Deviation (sigma): 10 Shape: 8 (7 for channels and one for the trend)

Reason: Similar to the intercept, the priors for the regression coefficients are set with a mean of 0 and a large standard deviation. This reflects no strong prior belief about the influence of any channel or the trend, allowing the data to inform the posterior distributions.

Standard Deviation of the Error Term (sigma):

Standard Deviation (sigma): 10

Reason: A HalfNormal distribution is used because the standard deviation must be non-negative. The large sigma ensures a non-informative prior, allowing the model to learn the appropriate scale of the residuals from the data.

3: Prior vs. Posterior Predictive Distributions

Prior Predictive Distribution Summary:

Mean: -0.289

Standard Deviation (SD): 32.312

HDI 3%: -60.614

HDI 97%: 61.059

MCSE (Mean Standard Error): 0.741

ESS (Effective Sample Size): 1845.0

R_hat: 1.2

Posterior Predictive Distribution Summary:

Mean: 11.768

Standard Deviation (SD): 0.344

HDI 3%: 11.128

HDI 97%: 12.421

MCSE (Mean Standard Error): 0.001

ESS (Effective Sample Size): 72463.0

R_hat: 1.04

Prior:

The model's initial beliefs before seeing any data suggest a wide range of plausible values for the parameter (x). The mean is -0.289 with a large standard deviation of 32.312, indicating significant uncertainty.

The HDI spans from -60.614 to 61.059, illustrating the broad spectrum of potential values the model considers plausible.

The low ESS (1845.0) and slightly elevated R_hat (1.2) suggest that more data or sampling iterations might be needed for better convergence.

Posterior:

After incorporating observed data, the model's beliefs about x have shifted significantly. The mean is now 11.768 with a much narrower standard deviation of 0.344, indicating higher confidence in the estimate.

The HDI for the posterior distribution 11.128 to 12.421 is narrower than the prior, showing reduced uncertainty after data inclusion.

High ESS values (72463.0) and a R_hat of 1.04 suggest good convergence and reliability of the posterior estimates.

Conclusion:

Prior: The model starts with a wide range of potential values for x, reflecting uncertainty without data influence.

Posterior: After observing data, the model updates x to a narrower range with higher confidence, indicating the impact of data on refining the parameter estimate.

These summaries illustrate how Bayesian inference progresses from initial uncertainty prior to more precise estimates posterior after considering observed data.

4: How good is your model performing? How you do measure it?

R-squared value of 0.453 indicates that the model explains approximately 45.3% of the variance in the observed revenue data, R-squared measures the proportion of the variance in the dependent variable revenue that is predictable from the independent variables model predictions. An R-squared of 0.453 means that the model accounts for 45.3% of the variability in revenue around its mean. The model captures a significant portion of the revenue variance, suggesting that the predictors adstocked spend and trend are meaningful in explaining revenue fluctuations.

A higher R-squared would indicate a better fit of the model to the data. However, the interpretation of what constitutes a "good" R-squared can vary based on the context and specific application.

5: What are your main insights in terms of channel performance/ effects?

Intercept:

The intercept is 11.768 with a narrow standard deviation sd = 0.024, indicating that on average, when all channel spends are zero, the predicted revenue is around 11.768.

Channel Effects (Beta Coefficients):

Channel 1 (beta[0]): Mean beta is -0.074 with a standard deviation of 0.043. This suggests that an increase in spend for Channel 1 is associated with a decrease in revenue, although the effect is uncertain (spanning from -0.155 to 0.006 in the HDI).

Channel 2 (beta[1]): Mean beta is -0.018 with a standard deviation of 0.033. Spending more on Channel 2 also shows a negative impact on revenue, but with less uncertainty (HDI from -0.077 to 0.047).

Channel 3 (beta[2]): Mean beta is 0.105 with a standard deviation of 0.031. This indicates a positive effect of spending on Channel 3 on revenue, with a relatively narrow HDI (from 0.048 to 0.166).

Channels 4 to 7 (beta[3] to beta[6]): Similar patterns are observed with varying degrees of impact and uncertainty.

Channel 7 (beta[7]): Mean beta is -0.120, indicating a negative impact of spending on Channel 7 on revenue, with an HDI from -0.207 to -0.040.

Model Precision and Convergence:

The model shows good convergence with all R_hat values close to 1.0 and effective sample sizes (ESS) generally above 2500, indicating reliable estimates.

Conclusion:

Channels vary in their effectiveness on revenue generation. Channels 3, 6, and 7 appear to have the most notable impacts, with Channel 3 showing the highest positive influence and Channel 7 the most significant negative influence.

Channels 1, 2, 4, and 5 exhibit more mixed effects, with some uncertainty in their impact on revenue.

6: Can you derive ROI (return on investment) estimates per channel? What is the best channel in terms of ROI?

ROI Estimates per Channel:

Channel 1:

Mean ROI: -8.067

95% HDI: [-17.551, 1.243]

Interpretation: On average, Channel 1 is estimated to have a negative ROI, indicating that the revenue generated does not sufficiently offset the spending. The HDI suggests high uncertainty, spanning from significantly negative to slightly positive ROI.

Channel 2:

Mean ROI: -7.140

95% HDI: [-33.148, 18.596]

Interpretation: Channel 2 also shows a negative mean ROI with a wide HDI, indicating substantial uncertainty in its impact on revenue relative to spending.

Channel 3:

Mean ROI: 0.734

95% HDI: [0.305, 1.167]

Interpretation: Channel 3 has a positive mean ROI, indicating that for every unit of currency spent, it returns a positive revenue. The HDI is relatively narrow, suggesting higher confidence in this estimate.

Channel 4:

Mean ROI: 0.742

95% HDI: [-0.330, 1.847]

Interpretation: Channel 4 also shows a positive mean ROI, with the HDI spanning from potentially negative to quite high ROI, reflecting some uncertainty.

Channel 5:

Mean ROI: -0.437

95% HDI: [-1.528, 0.642]

Interpretation: Channel 5 has a negative mean ROI, indicating that the revenue generated is less than the spending. The HDI suggests considerable uncertainty, with a potential for ROI ranging from strongly negative to slightly positive.

Channel 6:

Mean ROI: 1.374

95% HDI: [-0.332, 3.150]

Interpretation: Channel 6 shows the highest mean ROI among all channels, indicating that it generates the highest return relative to spending. However, the HDI is wide, indicating some uncertainty in the exact magnitude of ROI.

Channel 7:

Mean ROI: 0.459

95% HDI: [0.139, 0.791]

Interpretation: Channel 7 has a positive mean ROI, suggesting that it generates more revenue than the spending. The HDI is relatively narrow, indicating higher confidence in this estimate.

Best Channel in Terms of ROI:

Based on the mean ROI estimates, Channel 6 appears to be the best channel in terms of ROI, with a mean ROI of 1.374. This indicates that, on average, Channel 6 generates the highest return on investment relative to the other channels.