1. How do you model spend carry over?

Spend carryover refers to the delayed effect that marketing spend has on sales or revenue over time. Adstock transformation is a way to model these delayed effects. It works by applying a decay function to advertising spend, where the impact of previous spend decays over time but still carries forward to influence future sales. The adstocked spend is a weighted sum of current and past spends, with a decay rate controlling how much past spend contributes to future periods.

def apply\_adstock(spend, decay\_rate=0.5): adstock\_spend = np.zeros\_like(spend)

adstock\_spend[0] = spend[0] # initial spend value remains the same

for t in range(1, len(spend)):

adstock\_spend[t] = spend[t] + decay\_rate \* adstock\_spend[t - 1]

return adstock\_spend

A decay rate of 0.5 means that half of the previous week’s spend impact is carried over to the next week.

This transformed spend variable captures both immediate and delayed effects.

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

In Bayesian modeling, priors represent the initial beliefs or assumptions about the model's parameters before we see the data.

sigma = pm.HalfNormal("sigma", sigma=1.0)

intercept = pm.Normal("intercept", mu=0, sigma=10)

beta\_channel\_1 = pm.Normal("beta\_channel\_1", mu=0, sigma=sigma)

sigma (HalfNormal): The sigma parameter represents the noise or error term in the model. A HalfNormal prior with sigma=1.0 means that we believe the error in predicting revenue is generally small but can vary. Since sigma is positive by definition, a HalfNormal is used (it’s like a truncated Normal distribution).

intercept (Normal): The intercept term represents the baseline revenue when all other factors are zero. The prior for this is a Normal distribution with a mean of 0 and a large standard deviation (sigma=10), which reflects that we have little prior knowledge about the baseline revenue, so we allow for a wide range of possible values.

beta\_channel\_x (Normal): Each marketing channel’s effect (beta\_channel\_1, beta\_channel\_2, etc.) has a Normal prior centered around 0 (no effect) and a standard deviation controlled by sigma. This reflects the assumption that the impact of each channel on sales is uncertain, but we expect the average effect to be close to zero. If we had stronger prior knowledge of channel effects, we might adjust these priors accordingly.

trend, seasonality (Normal): The trend and seasonality components are modeled as Normal priors, which assume that any long-term trends (upward or downward) or seasonal effects on revenue are also uncertain but can be estimated from the data.

3. How are your model results based on prior sampling vs. posterior sampling?

Prior Sampling: In prior sampling, we generate samples from the model's priors before incorporating any observed data.

prior\_samples = pm.sample\_prior\_predictive(100)

The results from prior sampling are not yet informed by the data, and they represent what the model would predict before observing the actual data. These samples reflect only the priors you’ve set for the model parameters.

Posterior Sampling: In posterior sampling, you use the observed data to update the prior beliefs and generate samples from the posterior distribution. This represents the model's updated understanding of the parameters after observing the data.

trace = pm.sample(2000, tune=1000, target\_accept=0.9)

The posterior distribution incorporates both the prior beliefs and the data, so it provides more accurate estimates of the parameters (like the channel coefficients and the intercept).

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

R-hat values close to 1.0 indicate that the Markov chains have mixed well and that the model’s estimates are stable.

pm.summary(trace)

R-hat values should be close to 1 to ensure that the chains have converged. In the model all Rhat values are 1.

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

To understand the impact of each channel on revenue, you can look at the posterior distribution of the channel coefficients (e.g., beta\_channel\_1, beta\_channel\_2, etc.). These coefficients represent the incremental effect of each channel on revenue, after accounting for all other variables.

Positive coefficients: A positive value for a channel coefficient indicates that increased spend on that channel is associated with higher revenue.

Negative coefficients: A negative coefficient suggests that increasing spend on that channel could reduce revenue (e.g., due to diminishing returns or inefficiencies).

Magnitude of coefficients: The magnitude of the coefficient tells you the strength of the effect. Larger coefficients mean that the channel has a stronger impact on revenue.

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

Yes, ROI can be calculated for each channel using the estimated coefficients (beta\_channel\_x) and the corresponding spend.

The channel with the largest positive coefficient has the highest ROI. Among these, beta\_channel\_6 has the highest mean value at 1.791. This indicates that, on average, every unit of spend on channel\_6 is associated with an incremental increase of 1.791 units of revenue. Channel 6 has the highest ROI based on the estimated coefficient value of 1.791.