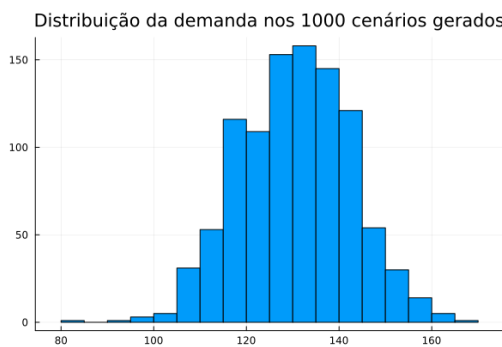


ML Under Modern Optimization Lens - From Predictive to Prescriptive Analysis - Exercícios

Giovanni Amorim

Junho, 2023

A lista tem como passo inicial a geração de dados de demanda e variáveis auxiliares por um processo dado. Na execução, foram gerados dados a partir de um processo ARMA(2,2), como especificado, porém a variável de demanda foi gerada de outra forma, a buscar assumir valores menores. Segue abaixo a distribuição da variável para as 1000 primeiras amostras geradas.



Solução do modelo de 2 estágios

A partir dos dados gerados, foi construído um modelo utilizando a lógica SAA (Sample Average Approximation), em que a decisão é tomada a partir da média entre cenários de custo dentro da amostra. Na rodada, a decisão tomada foi de compra de 138 unidades de jornal, assumindo um custo igual a $-1.875,56$.

Geração de novos 100 cenários

Foram gerados novos 100 cenários de variáveis auxiliares e demandas, seguindo a mesma lógica da amostra anterior.

Cálculo do custo em novo cenário considerando decisão fixa

Assumindo ainda a decisão fixa encontrada no primeiro item ($x = 138$), foi calculado o custo nos novos 100 cenários gerados, que foi igual a $-1.900,60$.

Solução do problema com KNN (k=5)

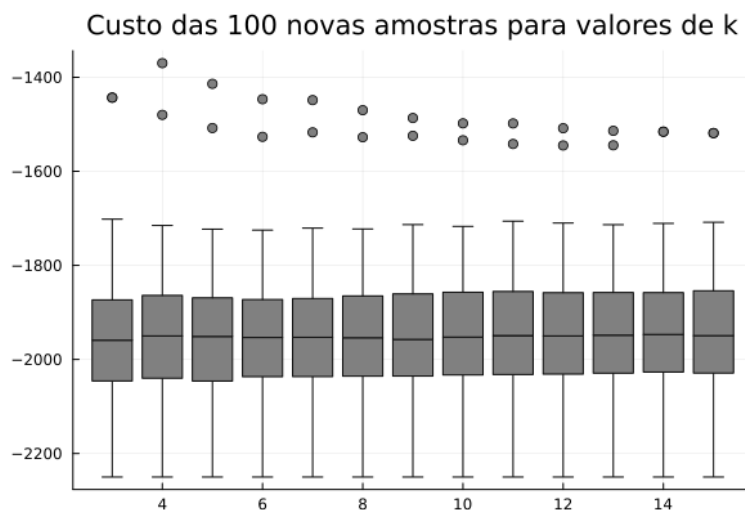
Para a solução utilizando o KNN, foi necessário seguir os seguintes passos:

1. Calcular uma árvore de proximidades baseada na primeira amostra, de 1000 cenários
2. Para cada amostra nos novos 100 cenários, encontrar o 5 cenários da amostra inicial mais próximos
3. Resolver o problema do jornaleiro considerando os 5 cenários mais próximos
4. Calcular o custo médio entre as amostras novas

Esse procedimento captura a mesma lógica do anterior, em que é utilizada a informação dos primeiros 1000 cenários na tomada de decisão em cenários novos. Nesse caso, o custo médio foi menor, de $-1.943,17$.

Variar k, calculando os custos

O gráfico a seguir mostra a distribuição dos custos para rodadas de solução com o KNN, com k variando:



Aqui podemos perceber que não há uma diferença grande entre valores de k próximos.

Cálculo de P para 200 novos cenários e modelo KNN (k=13)

Foi seguida a fórmula de P, considerando:

- $\hat{R}_v(\hat{z}_n)$: Custo alcançado pelo modelo KNN
- $\hat{R}_v(\hat{z}^{SAA})$: Custo alcançado pelo modelo SAA
- $\hat{R}_v(\hat{z}^*)$: Custo alcançado pela política ótima de informação completa, em que se compra exatamente a quantidade de jornais vendidos

A partir do cálculo:

$$P = \frac{\hat{R}_v(\hat{z}^{SAA}) - \hat{R}_v(\hat{z}_n)}{\hat{R}_v(\hat{z}^{SAA}) - \hat{R}_v^*(\hat{z}^*)}$$

chegou-se no valor aproximado de 0.6517, que indica a capacidade do modelo KNN, com K=13, de permitir uma tomada de decisão mais adequada à incerteza do processo.