# Retail Analytics

## Data-driven Newsvendor

Ben Lowery

STOR-i, Lancaster University

# Motivation

## Food Waste in Bakery Sector

- UK retail sector produced 240,000 tonnes of surplus food in 2015 (WRAP, 2017).
- Short shelf life of baked goods and excess food waste has far reaching impacts from profits to climate change.

## Safety Stock Planning

- Retailers have to keep in mind demand fluctuations and risk of stockout situations.
- Recent Example: Besins (menopause drug manufacturer) didn't anticipate spike in demand from increased media coverage.

- Setting: for a perishable product, over a single period with unknown demand ($D$), select order quantity ($Q$) that minimises cost and penalises over ($c_o$) and under ($c_u$) ordering.

**Newsvendor Model - Cost Minimisation**

$$\min_{Q \geq 0} \mathbb{E}\left[c_u (D - Q)^+ + c_o (Q - D)^+\right]$$

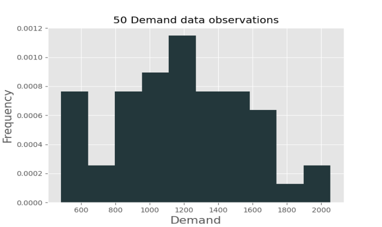
- Solution with known demand distribution and CDF $F$,

$$Q^*_{CM} = F^{-1}\left(\frac{c_u}{c_u + c_o}\right).$$

- Edgeworth (1888) establishment of *Inventory Theory* and then Arrow et al. (1951) formulation led to the Newsvendor Model.

# ISSUES AND SOLUTIONS

- In practical situations we do not know the demand distribution and rely on historically observed data.




- Many data-driven solutions have come about in recent years.
- Family of Distribution known? Parametric!
- Non-parametric approaches:
  - Sample Average Approximation (Levi et al., 2007).
  - Robust Optimisation (Scarf, 1957).
  - Feature based Linear Program (Beutel and Minner, 2012).

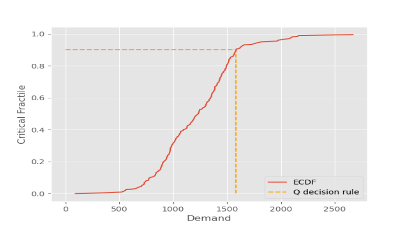
- Easy to implement and requires only historic demand observations.
- Assume each demand sample occurs with probability $\frac{1}{N}$.

### SAA - Cost Minimisation

$$\min_{Q \geq 0} \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}\left[c_u (d_i - Q)^+ + c_o (Q - d_i)^+\right].$$

- Decision rule comes from constructing empirical CDF.

$$\hat{Q}_{SAA} = \hat{F}^{-1}\left(\frac{c_u}{c_u + c_o}\right)$$



- Scarf (1957) proposed a distributionally robust optimisation model. Knowledge of first two moments only.
- Seek to minimise worst case expected costs by considering worst case distributions.

### Scarf's Min-max method

$$\min_{Q \geq 0} \quad \max_{D \in \mathcal{D}(\hat{\mu}, \hat{\sigma})} \mathbb{E}\left[c_u \left(D - Q\right)^+ + c_o \left(Q - D\right)^+\right]$$

- $\mathcal{D}(\mu, \sigma)$ is the set of distributions with mean $\mu$ and standard deviation $\sigma$.
- Optimal decision rule:

$$\hat{Q}_{SF} = \begin{cases} \hat{\mu} + \frac{\hat{\sigma}}{2}\left(\sqrt{\frac{c_u}{c_o}} - \sqrt{\frac{c_o}{c_u}}\right) & \text{if } 0 \leq \frac{c_u}{c_u + c_o} \leq \frac{\hat{\mu}}{\hat{\mu} + \hat{\sigma}}, \\ 0 & \text{Otherwise.} \end{cases}$$

# BUT WHAT ABOUT FEATURES?

- Focusing just on historic demand observations has lead to unsatisfactory analysis when forecasting demand (Beutel and Minner, 2012).
- Bakery sector for example: price, weather, seasonality may play a role.
- Feature-based Newsvendor models address this, LP approach:

$$
\begin{aligned}
\min_{\mathbf{q},\mathbf{v},\mathbf{s}} \quad & \sum_{i=1}^{n} (c_o v_i + c_u (d_i - s_i)) \\
\text{s.t.} \quad & v_i \geq \sum_{j=0}^{m} q_j X_{ji} - d_i, \quad i = 1, \ldots, n \\
& s_i \leq d_i, \quad i = 1, \ldots, n \\
& s_i \leq \sum_{j=0}^{m} q_j X_{ji} \quad i = 1, \ldots, n \\
& \mathbf{s}, \mathbf{v} \geq 0, \mathbf{q} \in \mathbb{R}.
\end{aligned}
$$

- $\mathbf{q}$ - decision variables.
- $v_i, s_i$ - inventory levels and satisfied demands.
- $d_i$ demand observations.
- $\mathbf{X}$ feature matrix.
- Optimal decision rule:

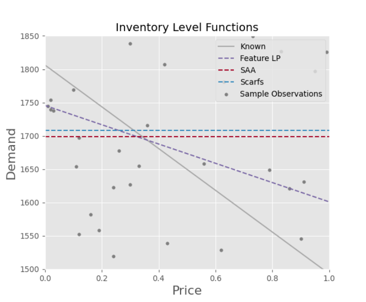
$$Q^*_{FBM} = q_0 + \sum_{j=1}^{m} q_j X_{j,0}.$$

# Featureless vs. Feature Based methods

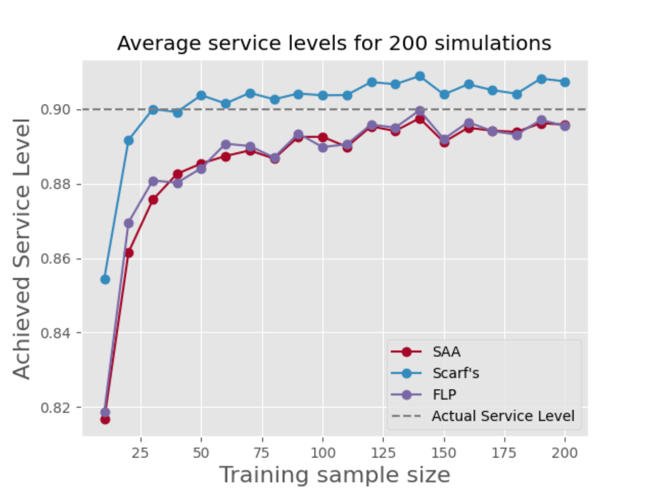
- Computational comparison between Feature-based method and Featureless (SAA and Scarf's min-max).
- Use artificially created data-set on one feature (price).

$$d_i = a - bp_i + \eta$$

- Metrics: achieved service level and average inventory levels.




# MORE RESULTS...



# Takeaways

- Scarf provides overly conservative results.
- Feature-based gets inventory levels very close to actual.
- No big difference between feature based and SAA on service levels.
- Future work:
    - Change underlying assumptions.
    - Increase number of features.

# Further Research Areas 1 - Machine Learning Models

- Seen a recent surge in Machine Learning based approaches (de Castro Moraes and Yuan, 2021).
- Issues with interpretability of Machine Learning models.
- However, Huber et al. (2019) concludes this may not matter if results and metrics are good enough.
- Potential future research in leveraging “Big Data” and develop better Feature based models.

# Further Research Areas 2

## Service Level Constraints

- Often retailers are bounded by "Service level agreements".
- New objective function:

$$\min_{Q \geq 0} \left\{ \mathbb{E}\left[(Q-D)^+\right] : \mathbb{P}[Q \geq D] \geq 1-\alpha \right\}.$$

- Future research in extending existent methodology to this problem setting.

## Censored Demand

- Frequent stockouts might lead to not knowing true demand and lost sales.
- Parametric methods from Conrad (1976) and Nahmias (1994).
- (non-parametric) data-driven approaches by Huh et al. (2011) and Sachs and Minner (2014).
- Future and current interest in extending to multi-product and multi-period settings.

# Thanks for listening!

# References

Arrow, K. J., Harris, T., and Marschak, J. (1951). Optimal inventory policy. *Econometrica*, 19(3):250–272.

Beutel, A.-L. and Minner, S. (2012). Safety stock planning under causal demand forecasting. *International Journal of Production Economics*, 140(2):637–645. Sixteenth internationalworkingseminaronproductioneconomics, Innsbruck, 2010.

Conrad, S. A. (1976). Sales data and the estimation of demand. *Journal of the Operational Research Society*, 27(1):123–127.

de Castro Moraes, T. and Yuan, X.-M. (2021). Data-driven solutions for the newsvendor problem: A systematic literature review. In *Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems*, pages 149–158. Springer International Publishing.

Edgeworth, F. Y. (1888). The mathematical theory of banking. *Journal of the Royal Statistical Society*, 51(1):113–127.

Huber, J., Müller, S., Fleischmann, M., and Stuckenschmidt, H. (2019). A data-driven newsvendor problem: From data to decision. *European Journal of Operational Research*, 278(3):904–915.

Huh, W. T., Levi, R., Rusmevichientong, P., and Orlin, J. B. (2011). Adaptive data-driven inventory control with censored demand based on kaplan-meier estimator. *Operations Research*, 59(4):929–941.

Levi, R., Roundy, R. O., and Shmoys, D. B. (2007). Provably near-optimal sampling-based policies for stochastic inventory control models. *Mathematics of Operations Research*, 32(4):821–839.

Nahmias, S. (1994). Demand estimation in lost sales inventory systems. *Naval Research Logistics (NRL)*, 41(6):739–757.

Sachs, A.-L. and Minner, S. (2014). The data-driven newsvendor with censored demand observations. *International Journal of Production Economics*, 149:28–36. The Economics of Industrial Production.

Scarf, H. E. (1957). *A Min-Max Solution of an Inventory Problem*, pages 201–209. RAND Corporation, Santa Monica, CA.

WRAP (2017). Quantification of food surplus, waste and related materials in the grocery supply chain. Technical report, WRAP.