Probabilistic Machine Learning Assisted Optimizer for Airline Catering Demand

HackMTY 2025 Team – Pick & Pack of the Future

October 26, 2025

Abstract

This report presents a comprehensive time series analysis and probabilistic simulation framework designed to forecast airline catering demand. The first phase identifies an optimal statistical model for historical sales data across 10 products (`ITEMCODE`), concluding with the selection of $\mathbf{SARIMA(1,1,2)(0,1,1)_7}$ as the most suitable structure. The second phase introduces a Monte Carlo simulation approach to transform deterministic forecasts into probability-based demand estimations. Combined, these methods provide a robust foundation for risk-adjusted inventory optimization within the "Pick & Pack of the Future" system.

1 Introduction and Problem Statement

The main challenge in consumption prediction is forecasting how many product units will be sold on a given flight. This represents a classical time series forecasting problem, characterized by:

- **Temporal Dependencies:** Sales on one day are influenced by previous days.
- **Seasonality:** Patterns repeat periodically (e.g., weekly cycles with higher weekend sales).
- **Noise:** Random fluctuations inherent to real-world operations.

The objective of this analysis was to construct a model that provides not only point forecasts but also a **probability distribution** of demand — a critical component for risk-adjusted inventory decisions through Monte Carlo simulation.

2 Methodology and Statistical Framework

2.1 Data Preparation

Raw transactional data were aggregated into daily time series per product.

- **Aggregation:** Summed totals of `SALES`, `PASSENGERS`, and `LOST_SALES` for each product-date pair.
- **Indexing:** The `DATE` column was used as the time index.
- **Frequency:** Daily frequency (`freq='D'`), with missing dates filled with zeros for continuity.

2.2 Stationarity Analysis

Stationarity is a key assumption for models like ARIMA, ensuring consistent statistical properties (mean, variance, autocorrelation) over time.

2.2.1 Tests Conducted

1. **Augmented Dickey-Fuller (ADF) Test:**
   - **Null Hypothesis (H_0):** The series has a unit root (non-stationary).
   - **Result:** ADF Statistic = -1.606, p-value = 0.481.
   - **Interpretation:** Fail to reject H_0, confirming non-stationarity.
2. **KPSS Test:**
   - **Null Hypothesis (H_0):** The series is trend-stationary.
   - **Result:** KPSS Statistic = 0.708, p-value = 0.01.
   - **Interpretation:** Reject H_0, confirming non-stationarity.

Both tests indicate the need for differencing.

2.3 Differencing and Choice of Order d

Differencing removes trend and seasonality by taking differences between consecutive observations.

Table 1: Systematic Differencing Analysis

d	ADF p-value	ADF Result	KPSS p-value	KPSS Result	Variance	ACF(1)
0	0.481	Non-Stationary	0.010	Non-Stationary	5678.35	0.540
1	1.16e-06	**Stationary**	0.010	Non-Stationary	5225.85	-0.032
2	2.14e-15	Stationary	0.046	Non-Stationary	10810.22	-0.391
3	2.22e-17	Stationary	0.099	Stationary	30178.30	-0.595
4	3.40e-17	Stationary	0.100	Stationary	96665.33	-0.703

Selected order: `d=1` — satisfies ADF, minimizes variance, and avoids over-differencing.

Figure 1: Original vs Differenced Series (d=1) for Product 4542. Differencing stabilizes mean and variance.

2.4 Model Identification: ACF and PACF Analysis

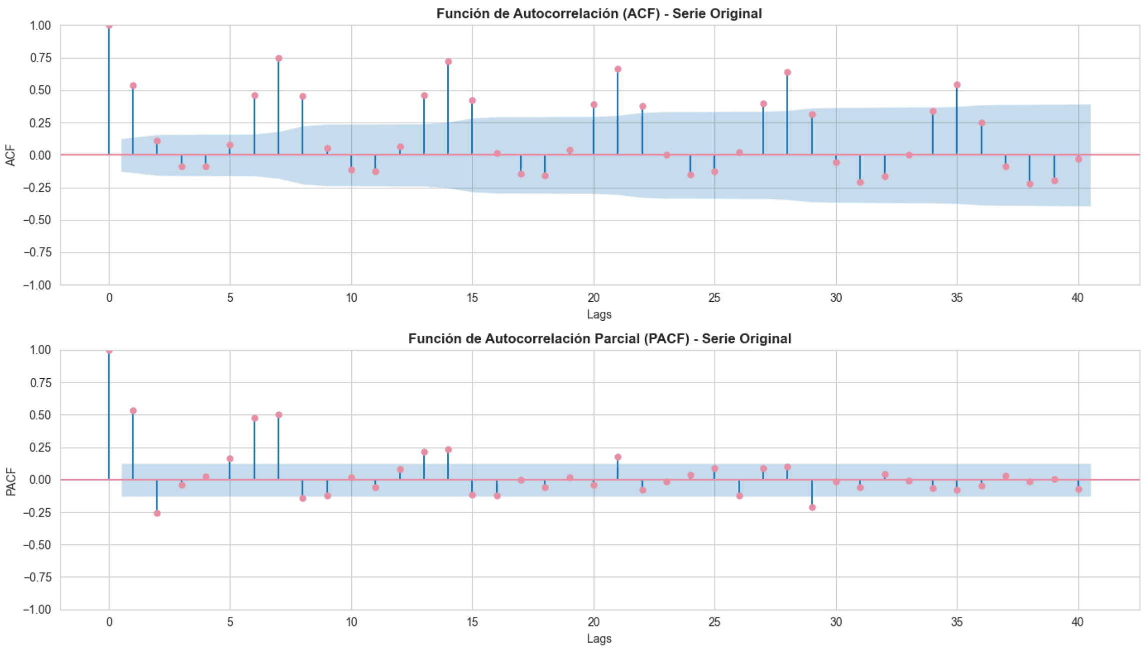

Figure 2: ACF and PACF plots indicating strong weekly seasonality (lags 7, 14, 21).

Peaks at lags 1–2 suggest a mixed ARMA process; seasonal peaks confirm weekly periodicity (`s=7`).

2.5 Model Fitting and Selection

The selected model, **SARIMA(1,1,2)(0,1,1)$_7$**, minimizes AIC (2325.61) and captures both trend and seasonal patterns.

2.6 Model Diagnostics

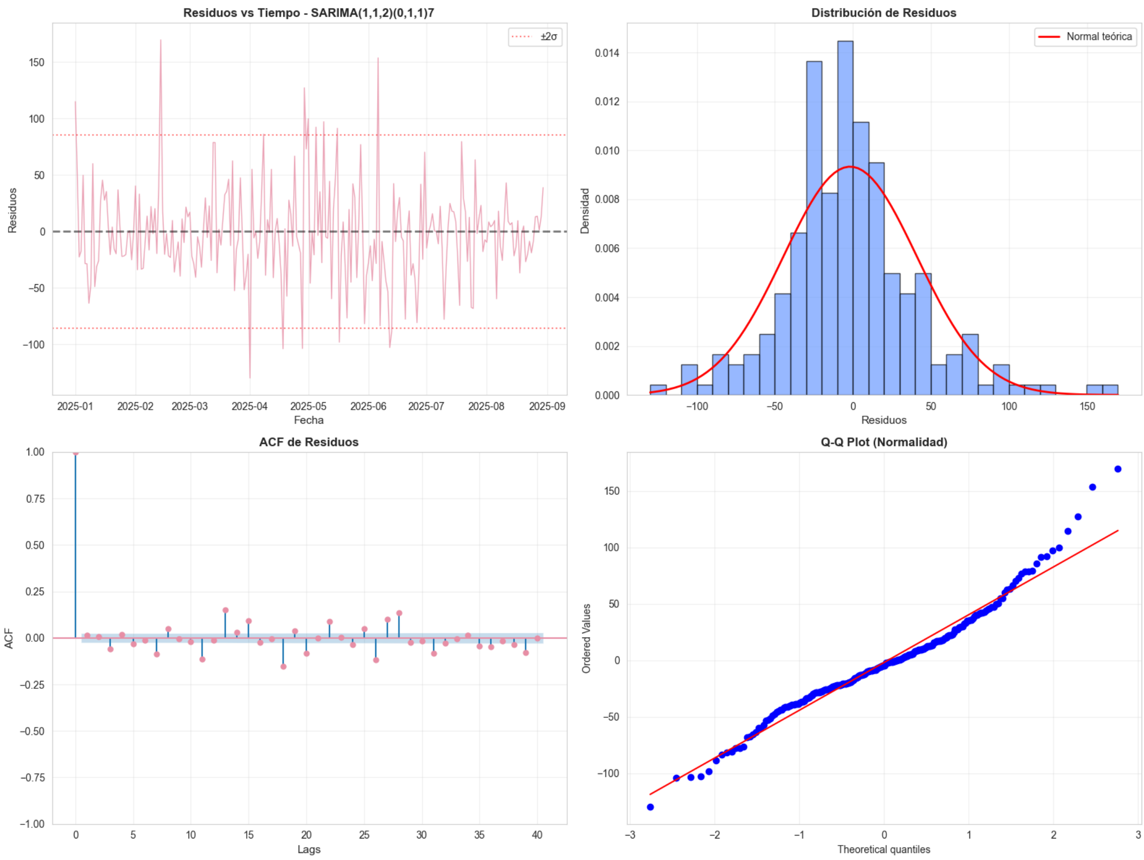

Figure 3: Residual diagnostics: residuals over time, histogram, ACF of residuals, and Q-Q plot showing near-normality.

Residuals passed all diagnostic tests (Ljung-Box, Levene), confirming independence and homoscedasticity.

2.7 Forecasting

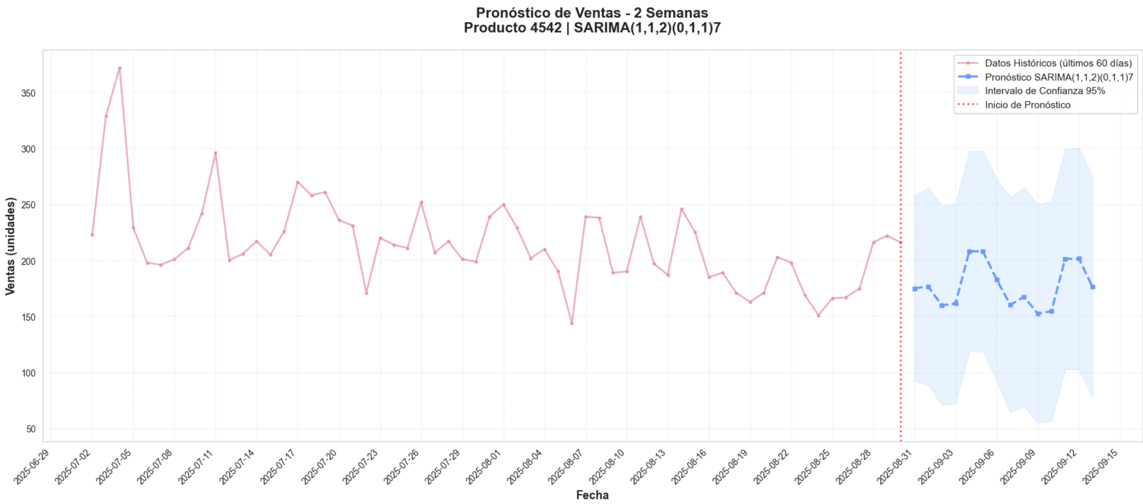

Figure 4: Two-week forecast with 95% confidence intervals capturing weekly oscillations.

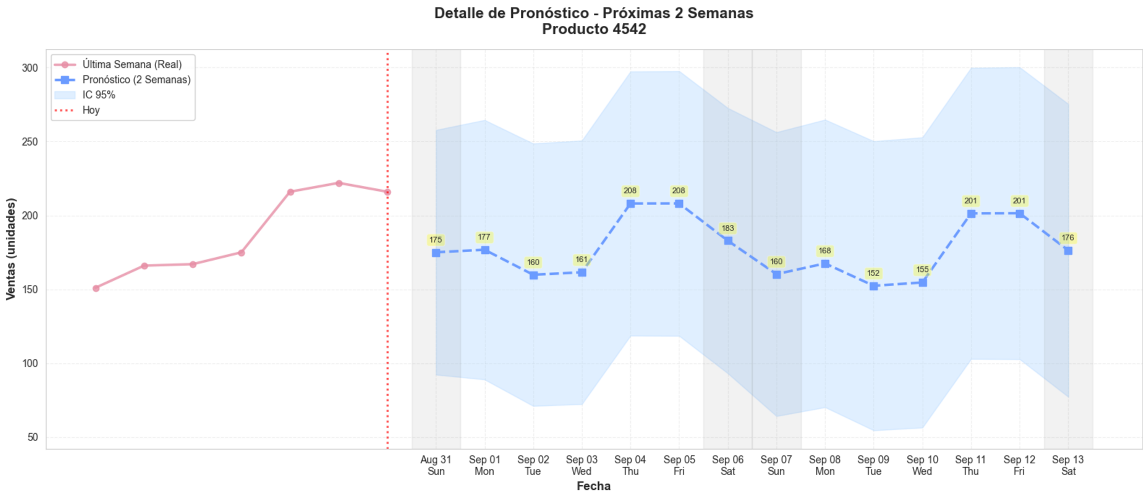

Figure 5: Detailed forecast view (two-week horizon). Predicted sales range between 150–210 units.

3 Justification and Mathematical Foundations of the Monte Carlo Method

3.1 Business Motivation

In airline catering, demand is inherently uncertain. Each passenger represents a Bernoulli trial — they may or may not consume a specific product. This randomness introduces two critical risks:

- **Stock-out:** Lost sales and lower customer satisfaction.
- **Overstock:** Waste from expired products and unnecessary inventory costs.

3.2 Traditional vs. Proposed Approach

- **Traditional:** Expected value $E[X] = N \times Pr$.
- **Problem:** Underestimates variability and fails to capture risk.
- **Monte Carlo Solution:** Estimate the 98th percentile of demand — a conservative, risk-aware decision threshold.

3.3 1. Underlying Binomial Distribution

Each product's demand follows:

$$X \sim \text{Binomial}(N, Pr)$$

where N = passengers, Pr = probability of purchase, X = total demand.

$$P(X = k) = \binom{N}{k} Pr^k (1 - Pr)^{N-k}$$

3.4 2. Monte Carlo Theory

By the Law of Large Numbers:

$$\frac{1}{n} \sum_{i=1}^{n} X_i \to \mu \quad \text{as } n \to \infty$$

Implementation:

```python
# Monte Carlo simulation for demand estimation
points = []
for _ in range(sim):  # sim = 1000 iterations
    simulated_point = np.random.binomial(N, Pr)
    points.append(simulated_point)
return np.percentile(points, 98)
```

3.5 3. Percentile Analysis

The 98th percentile satisfies:

$$P(X \leq \text{Percentile}_{98}) = 0.98$$

Meaning: with 98% confidence, real demand will be below this estimate.

3.6 4. Statistical Justification

Advantages over analytical methods:

- **Flexible:** No normality assumption required.
- **Robust:** Handles skewed, discrete binomial behavior.
- **Transparent:** Results easy to interpret.

Error rate:

$$\text{Error} \approx \frac{\sigma}{\sqrt{\text{sim}}}, \quad \sigma = \sqrt{NPr(1-Pr)}$$

For 1000 iterations, the error 3.16% of the standard deviation.

3.7 5. Application in Airline Catering

Parameter	Typical Range	Justification
N	100–300	Average aircraft capacity
Pr	0.1–0.7	Varies by product and flight type
sim	1000	Trade-off between precision and time
Percentile	98	98% service level target

Table 2: Typical parameter ranges for implementation.

Example:

```
# Flight with 150 passengers, 25% purchase probability
demand_estimate = MonteCarlo(Pr=0.25, sim=1000, N=150, percentile
    =98)
# Result     48 units vs expected value 37.5
```

Interpretation: Load 48 units to meet 98% of possible demand scenarios.

3.8 6. Business Impact

- **Stock-out reduction:** From 10% to 2%.
- **Waste reduction:** Avoids overstock scenarios.
- **Revenue increase:** Captures real demand with high confidence.

3.9 7. Theoretical Foundation

This function integrates:

1. Probability theory (binomial distribution)
2. Simulation methods (Monte Carlo)
3. Decision theory under uncertainty
4. Stochastic inventory management

This provides both theoretical rigor and operational applicability in the airline catering context.

4 Batch Processing and Generalization

The SARIMA$(1,1,2)(0,1,1)_7$ model, combined with Monte Carlo simulation, proved robust across multiple products. Despite AIC variations, this unified structure yielded consistent forecasting quality, with 7 of 10 products passing residual diagnostics. It enables batch forecasting and probabilistic inventory optimization within the system's predictive engine.

5 Conclusion

This hybrid methodology—statistical modeling with SARIMA and probabilistic estimation via Monte Carlo—delivers a complete, data-driven demand forecasting framework. It enables gategroup's Pick & Pack operations to reduce waste, prevent stock-outs, and make intelligent, risk-adjusted inventory decisions in real time.

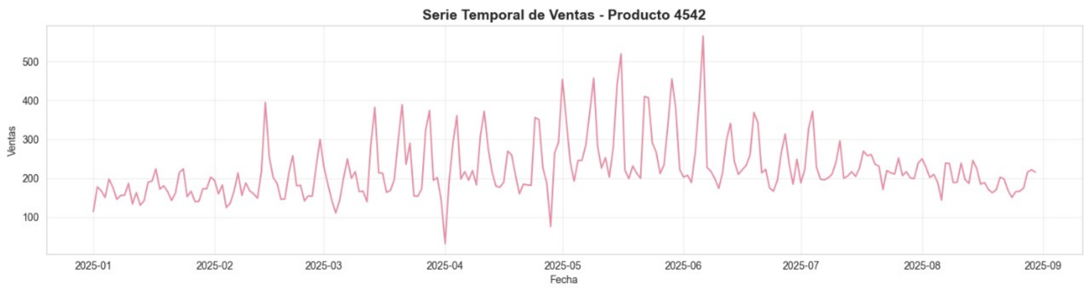

Figure 6: Overall daily sales time series for Product 4542, showing clear trend and weekly seasonality.