Mathematical Formulation of a Multi-Supplier Perishable Inventory MDP
with Stochastic Demand, Lead Times, and Spoilage Dynamics

Prepared for Simulation and Decision-Making Framework Development

Contents

1   Introduction

This report presents a full mathematical formulation of a Markov Decision Process (MDP) designed to manage a perishable pharmaceutical inventory system with:

- Multiple suppliers with heterogeneous lead times and prices.
- Inventory tracked by expiry using FIFO buckets.
- Pipelines for incoming orders and scheduled supply.
- Stochastic demand driven by seasonal and external factors.
- Inventory safety thresholds from demand forecasts.
- Actions that choose how much to order from each supplier (or hold off).
- A reward/cost system balancing shortages, spoilage, cost, and service reliability.

The model is suitable for Egypt's pharmaceutical import and hospital procurement context.

2   Time, Sets, and Indices

Time is discrete: $t = 0, 1, 2, \ldots$

- Shelf-life: N periods.
- Suppliers: $\mathcal{S} = \{1, \ldots, S\}$.
- Supplier s has lead-time bucket count L_s.

3   State Variables

At the start of period t, the system state is:

$$\mathsf{X}_t = \left(\mathbf{I}_t, \{\mathbf{P}_t^{(s)}\}_{s \in \mathcal{S}}, B_t, z_t \right)$$

Notation: We denote the state as X_t (not S_t) to maintain clarity that $\mathcal{S}$ is reserved exclusively for the set of suppliers. This prevents confusion between state variables and supplier index notation.

1. On-Hand Inventory by Expiry Buckets

$$\mathbf{I}_t = \left(I_t^{(1)}, I_t^{(2)}, \ldots, I_t^{(N)} \right)$$

where $I_t^{(1)}$ expires after this period and $I_t^{(N)}$ is freshly received.

2. Supplier Pipelines

For supplier s:

$$\mathbf{P}_t^{(s)} = \left(P_t^{(s,1)}, \ldots, P_t^{(s,L_s)} \right)$$

where $P_t^{(s,1)}$ arrives at the start of period t.

3. Scheduled/Committed Supply

Non-decision-based incoming supply is tracked with:

$$\widetilde{\mathbf{P}}_t^{(s)}$$

4. Backorders

$$B_t \geq 0$$

(If using a lost-sales model, set $B_t = 0$ always.)

5. Exogenous Variables

$$z_t$$

drives seasonal and contextual demand behavior.

4 Action Space

The decision action at time t:

$$\mathbf{a}_t = \left(a_t^{(1)}, a_t^{(2)}, \ldots, a_t^{(S)} \right)$$

Constraints may include:

- Supplier capacity: $a_t^{(s)} \leq U_s$

- MOQs: $a_t^{(s)} \in M_s \mathbb{Z}_{\geq 0}$

- Fixed ordering cost via binary trigger $y_t^{(s)}$

Holding off is represented by $\mathbf{a}_t = \mathbf{0}$.

5 Demand Process

Demand:

$$D_t \sim F(\cdot \mid z_t)$$

Examples:

- Poisson / Negative Binomial with time-varying mean.

- State-space forecast models.

Exogenous process evolves via:

$$z_{t+1} = g(z_t, \varepsilon_t)$$

6 Sequence of Events in Each Period

1. Arrivals

$$A_t = \sum_{s \in \mathcal{S}} \left(P_t^{(s,1)} + \widetilde{P}_t^{(s,1)} \right)$$

$$I_t^{(N)} \leftarrow I_t^{(N)} + A_t$$

2. Serve Demand FIFO

Let $R = D_t$. For $n = 1, \ldots, N$:

$$\text{take}_n = \min(I_t^{(n)}, R)$$

$$I_t^{(n)} \leftarrow I_t^{(n)} - \text{take}_n$$

$$R \leftarrow R - \text{take}_n$$

Boundary Condition: If $n = N$ and $R > 0$, the remaining demand R cannot be fulfilled from on-hand inventory and is recorded as backorders or lost sales.

Sales:

$$x_t = D_t - R$$

Backorders:

$$B_t^{\text{new}} = \max(R, 0)$$

Snapshot for cost:

$$\widehat{I}_t^{(n)} = I_t^{(n)}$$

3. Costs Before Aging

$$C_t^{\text{purchase}} = \sum_s \left(v_s a_t^{(s)} + K_s \mathbf{1}_{\{a_t^{(s)} > 0\}} \right)$$

$$C_t^{\text{hold}} = \sum_{n=1}^{N} h_n \widehat{I}_t^{(n)}$$

$$C_t^{\text{short}} = b \cdot B_t^{\text{new}}$$

$$c_t = C_t^{\text{purchase}} + C_t^{\text{hold}} + C_t^{\text{short}}$$

4. Aging and Spoilage

$$\text{Spoiled}_t = \widehat{I}_t^{(1)}$$

$$I_{t+1}^{(n)} = \widehat{I}_t^{(n+1)}, \quad n = 1, \ldots, N-1, \quad I_{t+1}^{(N)} = 0$$

Add spoilage cost:

$$c_t \leftarrow c_t + w \, \text{Spoiled}_t$$

5. Pipeline Shifts and New Orders

$$P_{t+1}^{(s,\ell)} = \begin{cases} P_t^{(s,\ell+1)}, & \ell < L_s \\ a_t^{(s)}, & \ell = L_s \end{cases}$$

$$\widetilde{P}_{t+1}^{(s,\ell)} = \widetilde{P}_t^{(s,\ell+1)}, \quad \widetilde{P}_{t+1}^{(s,L_s)} = 0$$

6. Backorder Update

$$B_{t+1} = B_t + B_t^{\text{new}}$$

7 Stochastic Lead Times

To model stochastic lead times while maintaining integer constraints, we introduce a Bernoulli random variable $\xi_t^{(s)} \in \{0, 1\}$ representing the advancement of the pipeline for supplier s at time t, governed by a transition probability matrix.

Formal Definition:

$$\xi_t^{(s)} \sim \text{Bernoulli}(p_s), \quad p_s \in [0, 1]$$

Alternatively, if using a state-dependent (Markovian) lead-time process with states i, j representing the current and next positions in the lead-time pipeline:

$$\xi_t^{(s)} \sim \text{Bernoulli}\left(\Pi_{i,j}^{(s)}\right)$$

where $\Pi^{(s)}$ is the state transition probability matrix for supplier s.

The pipeline evolution is defined as:

$$P_{t+1}^{(s,\ell)} = \begin{cases} P_t^{(s,\ell+1)} & \text{if } \xi_t^{(s)} = 1 \text{ (advance)} \\ P_t^{(s,\ell)} & \text{if } \xi_t^{(s)} = 0 \text{ (stall)} \end{cases}$$

with the boundary condition for the new order:

$$P_{t+1}^{(s,L_s)} = a_t^{(s)} \cdot \xi_t^{(s)} + P_t^{(s,L_s)} \cdot (1 - \xi_t^{(s)})$$

8 Survival-Adjusted Inventory Position

Let $\Phi_n(\cdot)$ be the Cumulative Distribution Function (CDF) of the sum of demand over n periods. Under FIFO issuance, the survival probability ρ_n for inventory in bucket n is determined by the probability that cumulative demand exceeds the inventory held in earlier buckets:

Corrected Formula:

$$\rho_n = P\left(\sum_{k=1}^{n} D_{t+k-1} > \sum_{j=1}^{n-1} I_t^{(j)}\right) = 1 - \Phi_n\left(\sum_{j=1}^{n-1} I_t^{(j)}\right)$$

Explanation: The sum includes buckets 1 through $n-1$ (inventory in earlier buckets), not bucket n itself. This ensures we correctly compute the probability that bucket n will be consumed before expiry, accounting for FIFO depletion of earlier inventory.

Let ρ_ℓ^{arr} = analogous probability for units arriving in ℓ periods.

$$\mathrm{IP}_t^{\mathrm{surv}} = \sum_{n=1}^{N} \rho_n I_t^{(n)} + \sum_{s=1}^{S} \sum_{\ell=1}^{L_s} \rho_\ell^{\mathrm{arr}} P_t^{(s,\ell)} - B_t$$

9 Safe Inventory Threshold

Forecast cumulative demand over horizon H: mean $\mu_{t:H}$, sd $\sigma_{t:H}$.

$$S_t^{\mathrm{safe}} = \mu_{t:H} + z_\alpha \sigma_{t:H}$$

Normality Assumption: This formula assumes that the cumulative demand $D_{t:t+H-1} = \sum_{k=0}^{H-1} D_{t+k}$ is approximately normally distributed. This is valid under the Central Limit Theorem when the number of periods H is sufficiently large and individual period demands are i.i.d. from a distribution with finite mean and variance. For non-normal demand (e.g., Poisson), this approximation may require $H \geq 5$–10 depending on the demand distribution's skewness.

Soft enforcement:

$$c_t \leftarrow c_t + \eta \cdot \max(S_t^{\mathrm{safe}} - \mathrm{IP}_t^{\mathrm{surv}}, 0)$$

Order-up-to target:

$$Q_t = \max(S_t^* - \mathrm{IP}_t^{\mathrm{surv}}, 0)$$

10 Tailored Base-Surge (Two Supplier Policy)

Supplier 1: slow/cheap (L_1, v_1)
Supplier 2: fast/expensive $(L_2 < L_1, v_2)$

Base level:

$$B = \mathbb{E}[D_{t:t+L_1-1}] + z_\alpha \cdot \mathrm{sd}(D_{t:t+L_1-1})$$

Slow supplier maintains base around B. Fast supplier surges:

$$a_t^{(2)} = \max(S_t^* - \mathrm{IP}_t^{\mathrm{surv}}, 0)$$

11   Rolling-Horizon MILP

The rolling-horizon controller solves the following finite-horizon optimization problem at each step t:

$$\min_{\mathbf{a}_\tau, \mathbf{y}_\tau} \sum_{\tau=t}^{t+H-1} \gamma^{\tau-t} \left(\sum_{s \in \mathcal{S}} \left(v_s a_\tau^{(s)} + K_s y_\tau^{(s)} \right) \right.$$
$$\left. + h \sum_{n=1}^{N} \widehat{I}_\tau^{(n)} + b \, \mathbb{E}[B_\tau^{\text{new}}] + w \, \mathbb{E}[\text{Spoiled}_\tau] \right)$$
$$+ \eta \sum_{\tau=t}^{t+H-1} \max(S_\tau^{\text{safe}} - \text{IP}_\tau^{\text{surv}}, 0)$$

Subject to the dynamic constraints (1)-(6) for all $\tau \in \{t, \ldots, t+H-1\}$.

Handling Stochasticity: The expectations $\mathbb{E}[B_\tau^{\text{new}}]$ and $\mathbb{E}[\text{Spoiled}_\tau]$ are computed using one of the following approaches:

1. **Sample Average Approximation (SAA):** Draw K scenarios of demand paths over the horizon and replace expectations with empirical averages.

2. **Deterministic Approximation:** Use forecast means μ_τ in place of demand randomness and compute deterministic spoilage/backorder predictions.

3. **Conditional Expectation:** Exploit Markovian structure to compute conditional expectations given current state.

Subject to pipeline balance and MOQ constraints:

$$0 \leq a^{(s)} \leq U_s y^{(s)}$$

12   Backorders vs Lost Sales

Backorders: carried forward and penalized. Lost sales: unmet demand disappears; penalties applied instead.

13   Bellman Equation

$$(\mathcal{T}V)(\mathsf{X}) = \min_{\mathbf{a} \in \mathcal{A}(\mathsf{X})} \left\{ c(\mathsf{X}, \mathbf{a}) + \gamma \, \mathbb{E}\left[V(\mathsf{X}') \mid \mathsf{X}, \mathbf{a} \right] \right\}$$

Transition Kernel: The expectation over the next state is defined explicitly as:

$$\mathbb{E}[V(\mathsf{X}') \mid \mathsf{X}, \mathbf{a}] = \sum_{\mathsf{X}'} P(\mathsf{X}' \mid \mathsf{X}, \mathbf{a}) \, V(\mathsf{X}')$$

where the transition kernel $P(\mathsf{X}' \mid \mathsf{X}, \mathbf{a})$ is determined by:

1. The stochastic demand distribution $F(\cdot \mid z_t)$.

2. The stochastic lead-time advancement if applicable (determined by $\xi_t^{(s)}$ or $\Pi^{(s)}$).

3. The exogenous process evolution $z_{t+1} = g(z_t, \varepsilon_t)$.

4. The deterministic inventory aging and pipeline dynamics.

Standard DP, ADP, or RL methods apply.

14 Theoretical Foundations and Structural Results

Base-Stock Policies

Proposition (Implicit): Under the following conditions:

1. Demand is i.i.d. across periods with finite mean and variance.

2. Holding and shortage costs are convex in inventory.

3. No fixed ordering costs or supplier switching costs.

4. Infinite horizon with discount factor $\gamma < 1$.

An optimal base-stock policy of the form $a_t^* = \max(S^* - \mathrm{IP}_t, 0)$ exists, where S^* is the order-up-to level. See Zipkin (2000), Porteus (2002) for proofs in the classical lost-sales and backorder settings.

Perishability Effects

With perishable inventory:

- The effective safety stock requirement increases due to potential spoilage.

- Optimal policies shift toward faster (more expensive) suppliers when inventory ages.

- The base-stock structure may be perturbed but remains approximately optimal if perishability is modeled via age-dependent holding costs h_n.

Multi-Supplier Structure

For two suppliers with different lead times and costs:

- The Tailored Base-Surge (TBS) policy allocates base demand to the slow supplier and surges to the fast supplier.

- Under certain conditions (e.g., cost and lead-time differences being sufficiently large), TBS is optimal or near-optimal.

- For multiple suppliers, greedy "cheapest effective arrival" rules provide good approximations.

To enhance computational efficiency and enable vectorized implementations, we formulate inventory aging as a matrix operation. Define the aging shift matrix:

$$\mathbf{A}_{\text{age}} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{N \times N}$$

Then the inventory evolution from aging can be written as:

$$\mathbf{I}_{t+1}^{\text{aged}} = \mathbf{A}_{\text{age}} \mathbf{I}_t$$

More generally, the full inventory update accounting for arrivals, consumption, and aging is:

$$\mathbf{I}_{t+1} = \mathbf{A}_{\text{age}} \left(\mathbf{I}_t - \mathbf{D}_t \right) + \mathbf{e}_N A_t$$

where:

- $\mathbf{D}_t$ is the demand vector (consumed from each bucket via FIFO).
- $\mathbf{e}_N = (0, \dots, 0, 1)^\top$ is the unit vector for the freshest bucket.
- A_t is the total arrivals at time t.

This matrix-based representation is suitable for GPU acceleration and batch optimization.

Policy Structure

- Base-stock or order-up-to policies emerge with convex costs. (See Scarf 1960, Zipkin 2000.)
- Multi-supplier allocation resembles "cheapest effective arrival" logic.
- With perishability, policy shifts toward faster suppliers when inventory is near expiry.
- Tailored Base-Surge becomes nearly optimal for two suppliers.

15   Notation and Parameter Index

This section summarizes all symbols, parameters, variables, and sets used throughout the model. Each item includes a brief explanation for clarity. All quantities are assumed to refer to a single SKU unless otherwise stated.

Sets and Indices

- $t = 0, 1, 2, \ldots$ Time periods.
- $\mathcal{S} = \{1, \ldots, S\}$ Set of suppliers.
- N Number of expiry buckets (shelf-life in periods).
- L_s Number of lead-time buckets for supplier s.

State Variables

- $I_t^{(n)}$ Inventory with n periods remaining before expiry ($n = 1$ expires soonest, $n = N$ freshest).
- $\mathbf{I}_t = (I_t^{(1)}, \ldots, I_t^{(N)})$ Entire on-hand inventory age profile.
- $P_t^{(s,\ell)}$ Pipeline quantity from supplier s arriving in ℓ periods.
- $\mathbf{P}_t^{(s)} = (P_t^{(s,1)}, \ldots, P_t^{(s,L_s)})$ Pipeline vector for supplier s.
- $\widetilde{P}_t^{(s,\ell)}$ Scheduled (non-decision) incoming supply for supplier s.
- B_t Backorders or unmet demand (0 if using lost-sales model).
- z_t Exogenous information state (seasonality, trend, demand drivers).

Decision Variables

- $a_t^{(s)} \geq 0$ Order quantity placed to supplier s at time t.
- $\mathbf{a}_t = (a_t^{(1)}, \ldots, a_t^{(S)})$ Full order vector.
- $y_t^{(s)} \in \{0, 1\}$ Ordering activation decision for supplier s (used with fixed order cost).

Demand and Stochastic Processes

- D_t Random demand in period t.
- $F(\cdot \mid z_t)$ Conditional demand distribution given state z_t.
- $z_{t+1} = g(z_t, \varepsilon_t)$ Exogenous process transition.
- $\Pi^{(s)}$ Transition matrix for stochastic lead times of supplier s.

Inventory Dynamics

- A_t Arrivals at time t from all pipelines.
- x_t Actual sales (fulfilled demand).
- B_t^{new} Newly created backorders at time t.
- Spoiled_t Inventory expiring at the end of period t.

Cost Parameters

- v_s Unit purchase cost from supplier s.

- K_s Fixed ordering cost for supplier s.

- h_n Holding cost per unit in expiry bucket n.

- b Shortage or backorder penalty cost per unit.

- w Spoilage (wastage) cost per expired unit.

- η Penalty weight for violating inventory safety threshold.

- γ Discount factor for infinite-horizon MDP.

Service and Forecast Parameters

- $\mu_{t:H}$ Forecasted mean cumulative demand over horizon H starting at t.

- $\sigma_{t:H}$ Standard deviation of cumulative demand over the horizon.

- z_α Normal quantile for service level α.

- S_t^{safe} Safety threshold for inventory at time t.

- S_t^* Order-up-to target inventory level.

Survival and Spoilage Modifiers

- $\rho_n \in [0,1]$ Probability that inventory with n remaining periods will be consumed before expiry.

- ρ_ℓ^{arr} Probability that a unit arriving in ℓ periods will be used before its expiry.

- $\text{IP}_t^{\text{surv}}$ Survival-adjusted inventory position.

Objective and Value Functions

- c_t Total cost incurred in period t.

- $r_t = -c_t$ Reward in reinforcement-learning formulation.

- $V(\mathsf{X})$ Value function for state X under optimal policy.

- $\mathcal{A}(\mathsf{X})$ Feasible action set given state X.

- $(\mathcal{T}V)(\mathsf{X})$ Bellman operator for discounted MDP.

Miscellaneous and Constraints

- U_s Maximum order size or capacity from supplier s.

- M_s Order lot size / MOQ (minimum order quantity).

- H Horizon length for safe-level computation or rolling horizon.