Problem Definition

In this model, the objective is to design a decision-making system for production planning under demand uncertainty. The products are categorized into several groups, and their demand is modeled through a set of possible scenarios. The model includes a set of products, product groups, and demand scenarios.

The key parameters of the model consist of the production capacity for each product, overarching sales targets, nominal and stochastic demand for each product, production cost, final profit, and the probability of each scenario occurring. The decision variables include surplus production quantity, sales quantity, surplus inventory level, inventory shortage, and the quantity of inventory substitution among similar products within a group.

This stochastic programming model aims to determine the optimal values for these variables in a way that, while considering production constraints and economic objectives, meets market needs under various demand scenarios, and simultaneously minimizes costs and the risk of inventory shortage or surplus.

Given that the demand for various products in this model is subject to uncertainty and each follows an independent probability distribution, accurately modeling the stochastic behavior of the entire system using a joint distribution function for all products is practically infeasible or computationally very complex. Under such circumstances, classical probabilistic modeling methods, particularly on a large scale, lack the necessary efficiency.

Therefore, one of the common and effective approaches for dealing with such complexities is the use of **scenario-based stochastic programming**. In this approach, instead of attempting to determine a continuous distribution for all random variables, a specific set of possible scenarios is generated, each representing a realization of the random variables (in this case, product demand). These scenarios are designed to cover a spectrum of plausible and probable future conditions, with each scenario assigned a specific probability of occurrence.

By incorporating these stochastic scenarios into the mathematical model, it becomes possible to analyze the system's behavior under uncertainty and make decisions that are statistically and economically robust against demand fluctuations. In what follows, the mathematical formulation of the stochastic programming model is presented, with the goal of optimizing production, allocation, and inventory decisions in the face of demand uncertainty.

Sets:	
I	Set of products
J	Set of product groups
SC	Set of demand scenarios
Parameters:	
Cap_i	production capacity of product i
MTP	macro target percentage
D_i	Nominal demand value of product i

D_i^{sc}	uncertain demand value of product I under scenario sc
$COGS_i$	Cost of goods sold
M_i	Marginal profit of product i
p_{sc}	The probability of occurrence of scenario i
Decision variables:	
S_i	Surplus production quantity of the product
A_i^{sc}	quantity of product i sold under scenario sc
U_i^{sc}	Excess inventory level of product *i* under scenario *sc*
L_i^{sc}	shortage quantity of product *i* under scenario *sc*
$O_{ii'}^{sc}$	inventory quantity of item *i* replaced by product *i'* of the same group under scenario *sc*

Stochastic mathematical model formulation:

$Max \quad Z = \sum_{i,sc} M_i . p_{sc} . A_i^{sc} - \sum_i (S_i + D_i) COGS_i$		1
$S_i + D_i \le Cap_i$	$\forall i$	2
$\sum_i S_i \le \sum_i D_i . MTP$	$\forall i$	3
$A_i^{sc} = Min(S_i + D_i, D_i^{sc})$	$\forall i, sc$	4
$L_i^{sc} = Max(0, S_i - D_i^{sc})$	$\forall i, sc$	5
$U_i^{sc} = Max(0, D_i^{sc} - S_i)$	$\forall i, sc$	6
$L_i^{sc} = \sum_{i' \in J_i} O_{i'i}^{sc}$	$\forall i, sc$	7
$U_i^{sc} = \sum_{i \in J_i} O_{ii'}^{sc}$	$\forall i, sc$	8
$A_i^{sc}, L_i^{sc}, U_i^{sc}, O_{ii'}^{sc}, S_i \ge 0$		9

Constraints 4 and 5 are nonlinear and must be linearized in order to obtain an exact optimal solution. The linearized equivalent of the above model is presented below:

$Max \quad Z = \sum_{i,sc} M_i . p_{sc} . A_i^{sc} - \sum_i (S_i + D_i) COGS_i$		1
$S_i + D_i \le Cap_i$	$\forall i$	2
$\sum_i S_i \le \sum_i D_i . MTP$	$\forall i$	3
$A_i^{sc} \le S_i + D_i$	$\forall i, sc$	4
$A_i^{sc} \le D_i^{sc}$		
$L_i^{sc} \ge S_i - D_i^{sc}$	$\forall i, sc$	5

$U_i^{sc} = D_i^{sc} - S_i$	$\forall i, sc$	6
$L_i^{sc} = \sum_{i' \in J_i} O_{i'i}^{sc}$	$\forall i, sc$	7
$U_i^{sc} = \sum_{i \in J_i} O_{ii'}^{sc}$	$\forall i, sc$	8
$A_i^{sc}, L_i^{sc}, U_i^{sc}, O_{ii'}^{sc}, S_i \geq 0$		9

Scenario Generation and Reduction Using the Sample Average Approximation (SAA) Method

Although the formulation of the presented stochastic model enables the analysis of decisions under uncertainty, the structural complexity arising from the stochastic nature of the problem remains a significant challenge. One of the main sources of this complexity is the presence of a large number of stochastic scenarios, which are generated to cover the possible range of future demand realizations. While these scenarios enhance the model's accuracy in representing reality, they exponentially increase its computational size, thereby imposing severe limitations on exact solutions in terms of time and computational resources.

In fact, using a small sample from the scenario space may lead to poor estimates of the objective function of the original model. On the other hand, excessive expansion of this space can substantially increase the model's solution time and computational burden, rendering even decomposition techniques impractical in real-world and large-scale problems. Therefore, there is a strong need for an efficient scenario reduction method that does not significantly compromise result quality.

One of the most effective approaches to addressing this challenge is the **Sample Average Approximation (SAA)** algorithm. This algorithm is based on the idea that, through random sampling of scenarios and statistical analysis of the obtained results, a reliable estimate of the objective function and optimal decisions can be achieved. The main steps of this algorithm are as follows:

Step 1: Sample Generation
Initially, M independent samples of the stochastic problem $(m = 1,2,, \dots, M)$ are generated, where each sample contains N independently and identically distributed (i.i.d.) random scenarios. Additionally, a larger reference set of scenarios with size $N' \gg N$ is generated to be used in subsequent steps for more accurate estimation.

Step 2: Solving Subproblems
Each of the m generated samples is solved using its corresponding scenario set, and the value of the objective function for each sample is recorded as v_m. Naturally, as the number of scenarios in each sample increases, the obtained solutions more closely approximate the true value of the objective function in the original model.

Step 3: Statistical Analysis of Results

The mean and variance of the objective function values obtained from solving the subproblems are calculated as follows:

$$\bar{v}_j^M = \frac{1}{M}\sum_{m=1}^{M} \hat{v}_j^M$$

$$S^2(\bar{v}_j^M) = \frac{1}{M(M-1)}\sum_{m=1}^{M} (\bar{v}_j^M - \hat{v}_j^M)^2$$

Step 4: Estimating the True Objective Function and Evaluating the Optimality Gap

By substituting the optimal solutions obtained from the initial samples into the larger reference set, a highly accurate estimate of the objective function for the original problem is achieved, denoted by $f_{N'}(\bar{v})$This value is considered an upper bound of the true objective function value. The variance of this estimate $S_{N'}(\bar{v})$ is also calculated, and the **optimality gap** of the model is determined as:

$$gap = \bar{v}_j^M - f_{N'}(x)$$
$$S_{gap}^2 = S_{N'}(\bar{v}) - S^2(\bar{v}_j^M)$$

If this gap falls within a statistically acceptable range, the solution can be regarded as a reliable estimate of the optimal solution for the stochastic model. Otherwise, the sampling and statistical analysis process is repeated until satisfactory convergence is achieved.

Overall, the SAA method demonstrates its primary advantage in balancing modeling accuracy under uncertainty with computational efficiency. In practical applications where the number of possible scenarios is extremely large, it is recognized as one of the most practical and effective techniques for reducing model complexity.

Pseudocode of the Sample Average Approximation Algorithm

1- Generation of 10,000 scenarios based on the six probability distribution functions provided for the demand of different product types
2- Selection of efficient scenarios and initial reduction of scenarios to 1,000 using the k-means approach, and storing them as the master scenarios N'
3- Determination of the initial problem parameters, sample size, scenario size, M, N.
4- Determination of the initial problem parameters, sample size, scenario size, and reference scenario size, respectively.
5-Start of the loop corresponding to the sample counter $m = 1,2,\dots,M$
    5-1- Selection of a number of random scenarios from the master set.
    5-2- Solving the Sample Average Approximation problem and determining the value of v_m

5-3- Solving the Sample Average Approximation problem and determining the value equal to the objective function value in the examined sample.

6- Fixing the first-stage variables.

7- Solving the Sample Average Approximation model with N' scenarios.

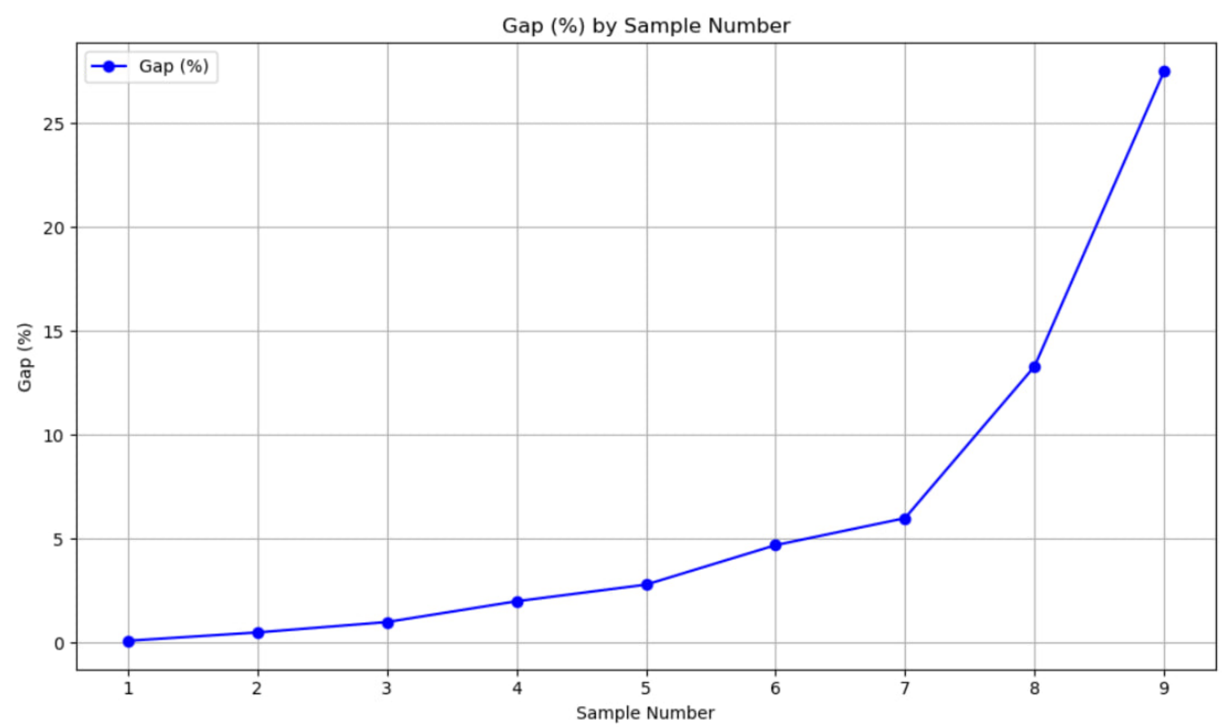
8- determining $f_{N'}(\bar{v})$ and $S_{N'}(\bar{v})$

9- Calculating the gap between the average of the sample solutions and the solution obtained from the large single-sample set.

10- End of the algorithm if the gap is less than the acceptable threshold or the number of iterations is sufficient; otherwise, update the values M، N and go to step 2.

The results obtained from solving the model with different scenarios and samples, along with the corresponding diagram, are presented as follows:

sample number	M	N	gap(%)
1	2	500	0.1
2	5	200	0.5
3	10	100	1
4	20	50	2
5	25	40	2.8
6	40	25	4.7
7	50	20	6
8	100	10	13.3
9	200	5	27.5



cap increase(%)	total marginal profit	gap	total surplus
5	2646460233	0.007	1978637.474
10	2697382327	0.012	2753223.027
15	2741201505	0.017	3445679.494
20	2779033448	0.022	4052261.894
25	2811815261	0.026	4585488.705
30	2840034671	0.021	4870658.496
35	2862540055	0.02	4882978.8
40	2862540055	0.02	4882978.8

The optimal value of the objective function, representing the expected profit margin, was obtained using the GAMS software as follows:

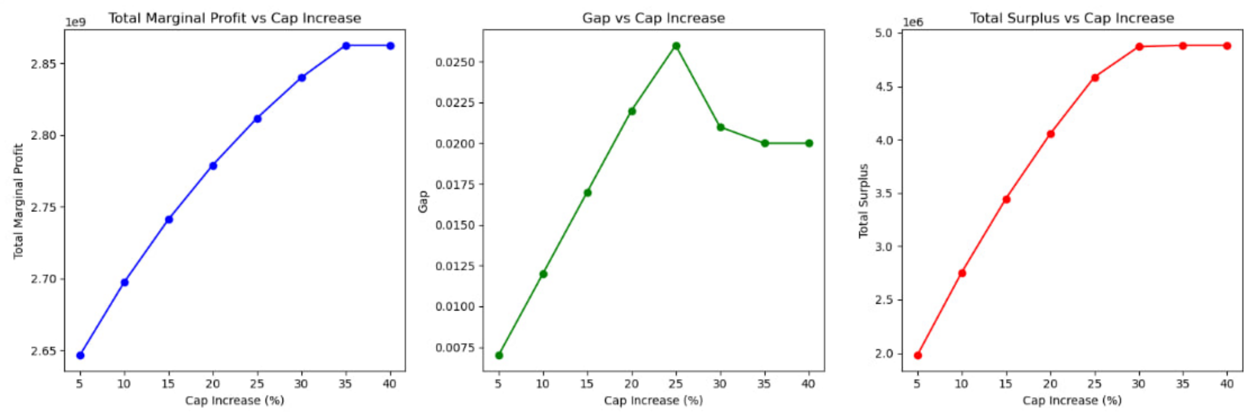
```
**** SOLVER STATUS     1 Normal Completion
**** MODEL STATUS      1 Optimal
**** OBJECTIVE VALUE       2587333713.4164
```

Sensitivity Analysis

In this section, a sensitivity analysis is conducted on the production capacity parameter. By incrementally increasing the capacity at a rate of 5%, the resulting changes in total production volume and the expected profit margin were examined. The corresponding variations are presented in the following tables and charts, illustrating the relationship between capacity expansion and its impact on key performance metrics of the model.

cap increase(%)	total marginal profit	gap	total surplus
5	2646460233	0.007	1978637.474
10	2697382327	0.012	2753223.027
15	2741201505	0.017	3445679.494
20	2779033448	0.022	4052261.894
25	2811815261	0.026	4585488.705
30	2840034671	0.021	4870658.496
35	2862540055	0.02	4882978.8

40	2862540055	0.02	4882978.8



As observed, increasing the production capacity up to 40% leads to a significant rise in both the expected profit margin and the optimal production volume. However, beyond this threshold, further capacity expansion has a diminishing effect, with negligible changes in the profit margin, indicating that the system reaches a saturation point where additional capacity no longer contributes meaningfully to overall performance.