Chapter 1: Introduction

Stochastic optimization is a framework for modelling optimization problems that involve uncertainty. On the other hand, Deterministic optimization problems are formulated with known parameters.

But real world problems almost invariably include some unknown and uncertain parameters.
Stochastic programming models take advantage of the fact that probability distributions governing the data are known or can be estimated.

The aim of stochastic programming is precisely to find an optimal decision in problems involving uncertain data. In this terminology, *stochastic* is opposed to *deterministic* and means that some data are random, whereas programming refers to the fact that various parts of the problem can be modeled as linear or nonlinear mathematical programs.

The field, also known as *optimization under uncertainty*, is developing rapidly with contributions from many disciplines such as operations research, economics, mathematics, probability, and statistics. The objective of this course is to provide a wide overview of stochastic programming.

Dealing With Randomness:

- For the case of Deterministic program, randomness is ignored, or it is dealt with by

i. Sensitivity analysis
– For large-scale problems, sensitivity analysis is useless

ii. "Careful" determination of instance parameters
– No matter how careful you are, you can't get rid of inherent randomness.

Soooooo
* Stochastic Programming is the way!

Types of Uncertainty:

Uncertainty appears everywhere (Uncertainty = Reality):

- Weather Related
- Demand in Supply chain
- Financial Uncertainty- Return in financial market
- Market Related Uncertainty
- Noisy, biased, incomplete data in machine learning
- Model uncertainty, state uncertainty in sequential decision making
- Competition
- Technology Related
- Acts of God

In an analysis of a decision, we would proceed through this list and identify those items that might interact with our decision *in a meaningful way!*

The Scenario Approach:

- A scenario-based approach is by no means the only approach to dealing with randomness, but it does seem to be a reasonable one.
- The scenario approach assumes that there are a finite number of decisions that nature can make (outcomes of randomness). Each of these possible decisions is called a scenario.

Example 1. Demand for a product is “low, medium, or high”.
Example 2. Weather is “dry” or “wet”.
Example 3. The market will go “up” or “down”

Even if the nature acts in a continuous manner, often a discrete approximation is useful.

A First Example:

- A Farmer can plant his land with either corn, wheat, or beans.
- For simplicity, assume that the season will either be wet or dry, – nothing in between.
- If it is wet, corn is the most profitable
- If it is dry, wheat is the most profitable.

Profit

	Corn	Wheat	Beans
Wet	100	70	80
Dry	-10	40	35

- So if the probability of a wet season is p, then the probability of dry season is $(1-p)$.

Then the expected profit of planting the different crops is:

Corn: $=100*p + (-10)* (1-p) = -10 + 110p$
Wheat: $= 70*p +40*(1-p) = 40 + 30p$
Beans: $80*p + 35*(1-p)= 35 + 45p$

What’s the Answer?

- Suppose $p = 0.5$, can anyone suggest a planting plan?

p = 0.5 , (1-p)= 1-0.5= 0.5

- Plant 1/2 corn, 1/2 wheat? Averaging the plantation

Expected Profit: $0.5 (-10 + 110(0.5)) + 0.5 (40 + 30(0.5)) = 50$

* Is this optimal?

No!

- Suppose $p = 0.5$, can anyone suggest another planting plan?
- Plant all beans!

Expected Profit: $35 + 45(0.5) = 57.5$!

* The expected profit in behaving optimally is = ((57.5-50)*100)/50 =15% better than in behaving reasonably.

***Averaging Solutions Doesn't Work!**

- The best decision for today, when faced with a number of different outcomes for the future, is in general not equal to the "average" of the decisions that would be best for each specific future outcome.
- That example is a little too simplistic for us to draw too many conclusions other than you cannot just average solutions.
- You *can't* replace random parameters by their mean value and solve the problem. This is (in general) not optimal either!

Mathematical form of Stochastic program (SP):

- Stochastic programs (linear, integer, Nonlinear) are mathematical programs where some of the data incorporated into the objective or constraints is uncertain.
- The outcomes are generally described in terms of elements x of a set X. Where X can be, for example, the set of possible demands over the next few months.

Deterministic Model: LP

Max / Min $\quad Z = c^T x$

subject to

$Ax \le B$

$x \ge 0$

(i) Uncertainty about the demand **b**,
(ii) uncertainty about the input prices **c,** and
(iii) uncertainty about the technical coefficient matrix **A** can be treated in SP.

These three types of uncertainties are related to parameters in SP.

Uncertainty about the demand **b**, uncertainty about the input prices **c,** and uncertainty about the technical coefficient matrix **A** can be treated in stochastic programming. These three types of uncertainties are related to parameters in stochastic programming.

An ordinary LP problem becomes stochastic when any or all elements of the set $\boldsymbol{\beta} = (\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c})$ of parameters depends on random set of natures $i.e. \beta = \beta(s), s \in W$, where W is the index set, *i.e.* the set of all possible states in the following form:

$$Minimize\ or\ Minimize\ \ z = cx$$

$$where\ x \in R = \{Ax \le b, x \ge 0\}$$

Scenario based Stochastic Models:

$$\text{Max / Min} \quad Z = c^T x$$

subject to

$$Ax \le B$$
$$T(\omega)\, x \le H(\omega)$$
$$x \ge 0$$

Stochastic program (SP): Stochastic programs (linear, integer, mixed integer, Nonlinear) are mathematical programs where some of the data incorporated into the objective or constraints is uncertain. Uncertainty is usually characterized by a probability distribution on the parameters. Although the uncertainty is rigorously defined, in practice it can range in detail from a few scenarios (possible outcomes of the data) to specific and precise joint probability distributions.

Classification of Stochastic Programs:

The stochastic programming problems can be classified as follows:

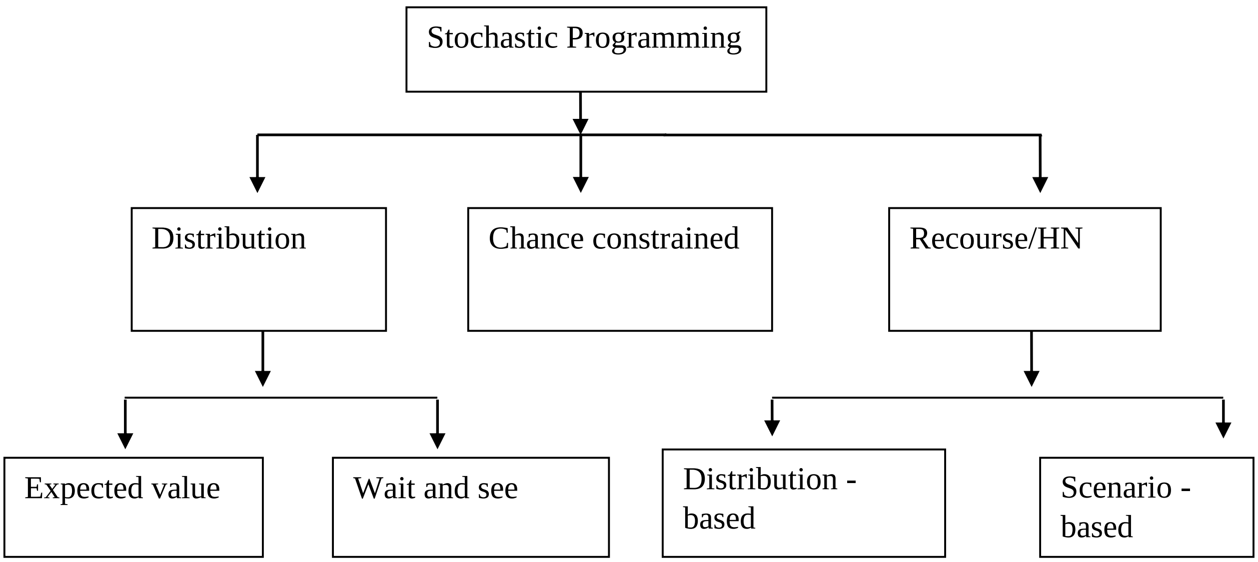

Figure 1: Taxonomy of SP problems

Distribution problem:

The optimization problems which provide the distribution of the objective function value for different realizations of the random parameters and also for the expected value of such parameters are broadly known as the distribution problems. It can be divided into two types of models. These are,

(i) Expected value problem (EVP).
(ii) Wait and see problem (WS).

(i) Expected value problem (EVP):

The expected value problem is constructed by replacing the random parameters by their expected values. Such an expected (EVP) model is thus an LP, as the uncertainty is dealt with before it is introduced in the under lying linear optimization model. It is common practice to formulate and solve the EVP problem in order to gain some insight into the decision problem.

(ii) Wait and see problem (WS) :

Wait and see problems assume that the decision maker is some-how able to wait until the uncertainty is resolved before implementing the optimal decisions. The approach therefore relies upon perfect information about the future. Because of its very assumptions such solutions cannot be implemented and is known as passive approach. Wait and see methods are often used to analyze probability distribution of the objective function value, and consists of a family of linear programming models each associated with an identical scenario.

Probabilistically constrained Model or Chance constrained Model:

In some cases, it may be more appropriate to try to find a decision which ensures that a set of constraints will hold with a certain probability.

A chance constrained optimization problem is a stochastic programming optimization problem involving one or more stochastic constrains that must hold simultaneously with a minimum given probability. It has the same LP objective function but instead of satisfying the deterministic constraints the objective function is subject to constraints that are satisfied in probabilistic manner. It has the following mathematical formulation:

$$Maximize\ z = \sum_{i=1}^{n} c_i x_i$$

Subject to

$$Probability\left(\sum_{i=1}^{n} a_{ij}\, x_i < b_j\right) \geq \sigma_i \qquad j = 1 \ldots m$$

$$x_i \geq 0\ , \quad i = 1..n$$

Where $0 \leq \sigma_i \leq 1$ for all $i = 1 \ldots n$. Here σ is used to control probability. Chance constrained optimization problems are challenging to solve in practice as they typically involve nonlinearities, non-convexities, discrete variables, and random variables. Even checking the feasibility of a given solution might already be hard in practice. Applications of chance constrained problems are numerous, e.g., management of resources, financial risk management, process engineering.

Assume that x is a vector of decisions. The general form of this problem is to

```
minimize                f(x1, x2, x3, ...., xn)

subject to     Pr[ g1(x1, x2, x3, ...., xn) <= 0
                                ...
                                ..
                   gm(x1, x2, x3, ...., xn) <= 0 ] >= alpha

                     h1(x1, x2, ... xn ) <= 0
                                ..
                     hk(x1, x2, ... xn ) <= 0

                 x1, x2, x3, ...., xn  in X
```

An example might be a delivery service that experiences random demands, and wishes to find the cheapest way to deliver its packages with a high probability.

Recourse problem (Here and now problem (HN)):

When dealing with "hare and now" decision problems, in general we do not have a necessarily penalize shortfall, but we might be able to take "corrective" action which is termed as recourse. For the case of a two stage problem, we make a decision now which is termed as **first stage decision**. Then nature makes a random decision called uncertainty. Finally we make a **second stage decision** that attempts to repair the uncertainty called recourse problem.

A two-stage HN model is presented below:

First stage

$$Min\ c_x x + c_z z + E[h(z, \xi)]$$
Subject to
$$A_x x + A_z z = b$$
$$0 \le x \le u_x,\ \ 0 \le z \le u_z$$

Second stage

$$h\ (z, \xi) = Min(fy)$$
Subject to
$$Tz + Wy = d$$
$$0 \le y \le u_y, \text{ where } \xi = vec(f, u_y, T, W, d)$$

Or

$$\begin{array}{ll} \min\limits_{x \in \mathbb{R}^n} & g(x) = c^T x + E[Q(x, \xi)] \\ \text{subject to} & Ax = b \\ & x \ge 0 \end{array}$$

where $Q(x, \xi)$ is the optimal value of the second-stage problem

$$\begin{array}{ll} \min\limits_{y \in \mathbb{R}^m} & q(\xi)^T y \\ \text{subject to} & T(\xi)x + W(\xi)y = h(\xi) \\ & y \ge 0 \end{array}$$

In the equations above the Wy term ensures that $Tx \le h$ remains feasible (seen by the fact that it depends on y, the decision variable of the second stage).

The first stage decisions are the vectors *x* and z, where the decisions *x* appear in the first stage but z links the first and second stage. The first stage decisions contribute directly to the objective value through the first stage cost. These decisions are constrained by linear constraints and simple upper bounds. The random variables affecting the second stage are described by the vectors ξ. In most general case the vector may affect all the parameters of the second stage problem. The cost of the second stage, $h\ (z, \xi)$, depends on the realizations ξ and on the decisions z. This cost is a random variable. The second stage problem is called recourse. The variable y is the second stage decision variable and it depends on *x*.

Alternatively,

Suppose x is a vector of decisions that we must take, and y(w) is a vector of decisions that represent new actions or consequences of x. Note that a different set of y's will be chosen for each possible outcome w.

The *Two-Stage* formulation is

```
minimize        f1(x) + Expected Value[ f2( y(w), w ) ]

subject to             g1(x) <= 0, ... gm(x) <= 0

                    h1(x, y(w) ) <= 0 for all w in W
                                  ..
                    hk(x, y(w) ) <= 0 for all w in W

                           x in X, y(w) in Y
```

The set of constraints h1 ... hk describe the links between the first stage decisions x and the second stage decisions y(w).

We require that each constraint hold with probability 1, or for each possible w in W. The functions f2 are quite frequently themselves the solutions of mathematical problems.
Or

$$
\begin{array}{lcccccccccccccc}
\text{Minimize} & f^T x & + & g^T y & + & p_1 h_1^T z_1 & + & p_2 h_2^T z_2 & + & \cdots & + & p_K h_K^T z_K & & \\
\text{subject to} & Tx & + & Uy & & & & & & & & & = & r \\
 & & & V_1 y & + & W_1 z_1 & & & & & & & = & s_1 \\
 & & & V_2 y & & & + & W_2 z_2 & & & & & = & s_2 \\
 & & & \vdots & & & & & & \ddots & & & & \vdots \\
 & & & V_K y & & & & & & & + & W_K z_K & = & s_K \\
 & x & , & y & , & z_1 & , & z_2 & , & \ldots & , & z_K & \geq & 0
\end{array}
$$

The deterministic equivalent problem can be solved using solvers such as CPLEX or GLPK, however it is important to note that if the number of scenarios is large, it may take a long time.

The *Multi-Stage* formulation:

Recourse models can be extended in a number of ways. One of the most common is to include more stages. With a multistage problem, we in effect make one decision now, wait for some uncertainty to be resolved (realized), and then make another decision based on what's happened. The objective is to minimize the expected costs of all decisions taken.

Scenario based problem:

The key to this approach is to create a set of scenarios. The idea is to construct or sample possible futures and solve the corresponding problems for these values. After having obtained a number of decisions in this way, we either pick the best of them, or we try to find to good combinations the decisions. If the problem for each scenario is a linear program, then the linear programs can be combined into one large program. The smaller linear programs are linked by variables corresponding to decisions that occur before it is known which scenario is active .one drawback of this approach is that the combined linear program can get very large if there are a lots of scenarios.

Scenario formulation:

There are several ways to formulate an SP. Here focus has been on formulations that explicitly represent the information process with in the sequence of decisions that are made.

An alternate, but equally valid, representation of the problem is one in which a problem is formulated for each possible scenario and constraints are added to ensure the information structure associated with the decision process honored. In this case, we begin by representing all decision variables as if they were permitted to depend on the specific scenario encountered, which leads to the scenario problems for each $\omega \in \Omega$:

$$\begin{aligned} &Minimize \qquad cx_\omega + g_\omega y_\omega \\ &Subject\ to \\ &\qquad T_\omega x_\omega + w_\omega y_\omega \geq r_\omega \\ &\qquad x_\omega, y_\omega \geq 0 \end{aligned}$$

Without the introduction of additional constraints, we obtain a situation in which $\{(x_\omega , y_\omega)\}_{\omega \epsilon \Omega}$ vary freely in response to each specific scenario, while others cannot. We can remedy this by including constraints that ensure that decision sequence honors the information structure present in the problem as follows:

$$Minimize \qquad \sum_{\omega \epsilon \Omega} (cx_\omega + g_\omega y_\omega) P_\omega \tag{1}$$

$$Subject\ to$$

$$T_\omega x_\omega + w_\omega y_\omega \geq r_\omega \tag{2}$$

$$x_\omega - x = 0 \tag{3}$$

$$x_\omega , y_\omega \geq 0$$

Here constraints such as (3) are known as non-anticipativity constraints and ensure that decisions honor the information structure of the problem. Note that in (3) we use a free variable, *x*, to constraint the scenario dependent, first stage variables $\{x_\omega\}_{\omega \epsilon \Omega}$ to be equal.

Applications of Stochastic Programming:

Stochastic programming has been applied to a wide variety of areas which are listed below.

- Manufacturing Production Planning
- Manufacturing production capacity planning
- Electrical generation capacity planning
- Machine Scheduling
- Freight scheduling
- Dairy Farm Expansion planning
- Macroeconomic modeling and planning
- Timber management
- Asset Liability Management
- Portfolio selection
- Traffic management
- Optimal truss design
- Automobile Dealership inventory management
- Lake level management

A Random Linear Program

$$\text{Min} \quad Z = x_1 + x_2$$
$$\text{subject to} \quad \omega_1 x_1 + x_2 \geq 7$$
$$\omega_2 x_1 + x_2 \geq 4$$
$$x_1, x_2 \geq 0\,.$$
$$\omega_1 \approx U[1,\, 4]$$
$$\omega_2 \approx U[1/3,\, 1]$$

What To Do?

- How do we solve this problem?
- What do you *mean* by solving this problem?
- Suppose it is possible to decide about *x* *after* the observation of the random vector *w*?

We can interpret this as a ***wait-and-see*** approach

* Can we solve the problem then?
It’s just a simple deterministic linear program!

Here and Now

* Generally, “wait-and-see” is not an appropriate model of how things work.
-We need to decide on *x* *before* knowing the values of *w*.
* In order for the problem to make sense in this case, we need to decide what to do about not knowing *w*1, *w*2.
*Three suggestions

Guess at uncertainty
Probabilistic Constraints
Penalize Shortfall

Guess Away!
* We will guess reasonable values for $w1$, $w2$
Three (reasonable) suggestions – each of which tells us something about our level of "risk"

***Unbiased*: Choose mean values for each random variable**
***Pessimistic*: Choose worst case values for w**
***Optimistic*: Choose best case values for w**

Unbiased (Average)

$$\hat{\omega} = E(\omega) = \left(\frac{5}{2}, \quad \frac{3}{2}\right) = \text{Average value}$$

$$\text{Min} \quad Z = x_1 + x_2$$

$$\text{subject to } \frac{5}{2}x_1 + x_2 \geq 7$$

$$\frac{3}{2}x_1 + x_2 \geq 4$$

$$x_1, x_2 \geq 0 .$$

$$\text{OF value} = z = \frac{50}{11}$$

$$(x_1, \quad x_2) = \left(\frac{18}{11}, \quad \frac{32}{11}\right)$$

Pessimistic

$$\hat{\omega} = E(\omega) = \left(1, \quad \frac{1}{3}\right)$$

$$\text{Min} \quad Z = x_1 + x_2$$

$$\text{subject to } x_1 + x_2 \geq 7$$

$$\frac{1}{3}x_1 + x_2 \geq 4$$

$$x_1, x_2 \geq 0 .$$

$$\text{OF value} = z = 7$$

$$(x_1, \quad x_2) = (0, \quad 7)$$

Optimistic

$$\hat{\omega} = E(\omega) = (4, \quad 1)$$

$$\text{Min} \qquad Z = x_1 + x_2$$

$$\text{subject to} \quad 4x_1 + x_2 \geq 7$$

$$x_1 + x_2 \geq 4$$

$$x_1, x_2 \geq 0 \, .$$

$$\text{OF value} = z = 4$$

$$(x_1, \quad x_2) = (4, \quad 0)$$

Chance Constrained

Another (probably more reasonable) approach. Let's enforce that the *probability* of a constraint holding is sufficiently large. Let's add the constraints

$$P\{\omega_1 x_1 + x_2 \geq 7\} \geq \alpha_1$$

$$P\{\omega_2 x_1 + x_2 \geq 4\} \geq \alpha_2$$

Or

$$P\{\omega_1 x_1 + x_2 \geq 7, \; \omega_2 x_1 + x_2 \geq 4\} \geq \alpha$$

Introduction

Operations research has been particularly successful in two areas of decision analysis:

(i) **Deterministic optimization**: optimization of problems involving many variables when the outcome of the decisions can be predicted with certainty, and

(ii) **Stochastic optimization**: the analysis of situations involving a few variables when the outcome of the decisions cannot be predicted with certainty.

Deterministic Solution:

The optimization algorithm determines the *best* solution given the parameters of the model. The modeling and decision analysis process is illustrated in Fig. 1 where the shape labeled *situation* represents the real problem under consideration. Various assumptions and abstractions are applied, including the assumption of deterministic information, to obtain a mathematical *model*. The model is input to a computer where an *algorithm* determines the optimal *decision*, represented by the vector **x**.

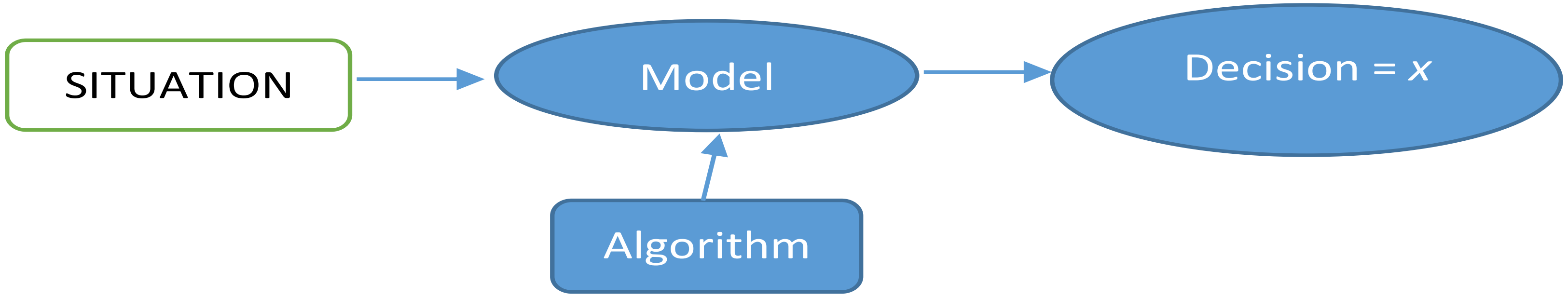

Figure 1: The deterministic approach to decision making

Deterministic Model: LP

$$\text{Max / Min} \qquad Z = c^T x$$

$$\text{subject to}$$

$$Ax \le B$$

$$x \ge 0$$

$$\text{Here } A = \begin{pmatrix} a_{11} & a_{12} & \ldots\ldots. & a_{1n} \\ a_{21} & a_{22} & \ldots.. & a_{2n} \\ . & . & . & . \\ . & . & . & . \\ a_{m1} & a_{m2} & \ldots\ldots. & a_{mn} \end{pmatrix} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ b_m \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix}$$

Disadvantage of Deterministic model:

- The fault with this approach is that the decision ***x*** is optimum for the model and not the situation.
- It is usually readily apparent to the manager with the task of implementing the decision, that ***x*** is not at all appropriate for application to the situation.
- The primary reason for this lies in the assumption of deterministic parameters.
- When the situation involves uncertainty or risk, the kind of decision taken is quite different than if it does not.
- Real decision makers hedge against various possible futures.

Example:

Imagine a company that provides energy to households.

The company is responsible for delivering energy to households based on their demand.

The problem could be solved as an LP with constraints based on demand from households.

But

future demand of households is not always known and is likely dependent on factors such as the weather and time of year.

Therefore, there is uncertainty and the basic LP model will not be sufficient.

Stochastic program (SP):

Stochastic programs (linear, integer, Nonlinear) are mathematical programs where some of the data incorporated into the objective function or constraints is uncertain.

Uncertainty is usually characterized by a probability distribution on the parameters.

Although the uncertainty is rigorously defined, in practice it can range in detail from a few scenarios to specific and precise joint probability distributions.

The outcomes are generally described in terms of elements x of a set X. Where X can be, for example, the set of possible demands over the next few months.

When some of the data is random, then solutions and the optimal objective value to the optimization problem are themselves random.

A distribution of optimal decisions is generally un implementable. Ideally, we would like one decision and one optimal objective value.

Scenario based Stochastic Models:

$$\text{Max / Min} \qquad Z = c^T x$$

subject to

$$Ax \le B$$
$$T(\omega)\, x \le H(\omega)$$
$$x \ge 0$$

Here $A = \begin{pmatrix} a_{11} & a_{12} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots.. & a_{2n} \\ . & . & . & . \\ . & . & . & . \\ a_{m1} & a_{m2} & \ldots\ldots & a_{mn} \end{pmatrix}$ $B = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ b_m \end{pmatrix}$, $T(\omega) = \begin{pmatrix} a_{11}(\omega) & a_{12}(\omega) & \ldots\ldots & a_{1n}(\omega) \\ a_{21}(\omega) & a_{22}(\omega) & \ldots.. & a_{2n}(\omega) \\ . & . & . & . \\ . & . & . & . \\ a_{l1}(\omega) & a_{l2}(\omega) & \ldots\ldots & a_{\ln}(\omega) \end{pmatrix}$, $x = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix}$

$$H(\omega) = \begin{pmatrix} h_1(\omega) \\ h_2(\omega) \\ . \\ . \\ h_l(\omega) \end{pmatrix}$$

This problem can be formulated according to:

$$\text{Max / Min} \qquad Z = c^T x + Q(x)$$

subject to

$$Ax \le B$$

$$T(\omega)\,x \le H(\omega)$$

$$x = \{x_1, x_{2,} \ldots . x_n\} \ge 0$$

Here $Q(x) = E_\omega Q(x, \varpi)$

$Q(x,\varpi) = \min \left| q(\varpi)^T y(\varpi) \right|$ where $W(\varpi) y(\varpi) \ge H(\omega) - T(\omega)x$ and ϖ is a scenario or possible outcome or random variable.

General Stochastic Programming

$$Min \quad c^T x + \sum_{s=1}^{S} p_s q_s^T y_s$$

$$s.t. \quad Ax \leq b$$

$$T_s\, x + W_s\, y_s = h_s$$

$$x,\, y_s \geq 0, \quad s = 1,2,\ldots., S$$

Max/ Min Z = Deterministic part + Stochastic part

s.t. Deterministic constraints
Stochastic constraints

Numerical Example

Suppose we have the following optimization problem:

$$\begin{array}{rl} \max & 5x_1 + 10x_2 - 4y_1 - 6y_2 \\ \text{subject to} & y_1 + y_2 = 10 \\ & x_1 + 3x_2 \leq 14 \\ & x_1, x_2 \geq 0 \end{array}$$

This is a simple linear optimization problem with optimal solution set $x_1 = 14, y_1 = 10, z = 30$.

Now assume that variables x_1 and x_2 are uncertain and that there are three different scenarios, $(1.5X, X, 0.7X)$ for the values of x_1 and x_2, each occurring with a probability of 1/3. This new problem involves uncertainty and is thus considered a stochastic problem. We must now partition x_1 and x_2 into $x_{1,1}, x_{1,2}, x_{1,3}$ and $x_{2,1}, x_{2,2}, x_{2,3}$ respectively.

3-Scenarios represents 3 independent deterministic LP Models

Scenario-1	Max $Z = (1.5*(5x_1 + 10x_2)) - 4y_1 - 6y_2$ subject to $y_1 + y_2 = 10,$ $1.5*(x_1 + 3x_2) \le 14$ $x_1, x_2, y_1, y_2 \ge 0$	OBJECTIVE FUNCTION VALUE 1) 30.00000 VARIABLE VALUE RC X1 9.333333 0.000000 X2 0.000000 7.500000 Y1 10.000000 0.000000 Y2 0.000000 2.000000
Scenario-2	Max $Z = (1*(5x_1 + 10x_2)) - 4y_1 - 6y_2$ subject to $y_1 + y_2 = 10,$ $1*(x_1 + 3x_2) \le 14$ $x_1, x_2, y_1, y_2 \ge 0$	OBJECTIVE FUNCTION VALUE 1) 30.00000 VARIABLE VALUE RC X1 14.0000 0.000 X2 0.000000 5.000000 Y1 10.000000 0.000000 Y2 0.000000 2.000000
Scenario-3	Max $Z = (0.7*(5x_1 + 10x_2)) - 4y_1 - 6y_2$ subject to $y_1 + y_2 = 10,$ $0.7*(x_1 + 3x_2) \le 14$ $x_1, x_2, y_1, y_2 \ge 0$	OBJECTIVE FUNCTION VALUE 1) 30.00000 VARIABLE VALUE RC X1 20.000000 0.000000 X2 0.000000 3.500000 Y1 10.000000 0.000000 Y2 0.000000 2.000000

Deterministic equivalent of the General stochastic LP problem

Max $$Z = \frac{1}{3}\left(1.5*\left(5x_{11} + 10x_{21}\right)\right) + \frac{1}{3}\left(1*\left(5x_{12} + 10x_{22}\right)\right) + \frac{1}{3}\left(0.7*\left(x_{13} + x_{23}\right)\right) - 4y_1 - 6y_2$$

subject to

$$y_1 + y_2 = 10,$$
$$1.5*\left(x_{11} + 3x_{21}\right) \le 14$$
$$1*\left(x_{12} + 3x_{22}\right) \le 14$$
$$0.7*\left(x_{13} + 3x_{23}\right) \le 14$$
$$x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23}, y_1, y_2 \ge 0$$

$$Min \quad c^T y + \sum_{s=1}^{3}\sum_{i=1}^{2} p_s q_s^T x_{is}$$
$$s.t. \quad y_1 + y_2 = 10$$
$$\sum_{i=1}^{2} q_s x_{is} \le 14, \quad s = 1,2,3$$
$$x, y_s \ge 0, \quad s = 1,2,\ldots, S$$

Once turned into the discrete version, the problem is reformulated as shown below and can be solved once again using linear programming.

$$
\begin{aligned}
\max \quad & \tfrac{1}{3}(7.5x_{1,1} + 15x_{2,1}) + \tfrac{1}{3}(5x_{1,2} + 10x_{2,2}) + \tfrac{1}{3}(3.5x_{1,3} + 7x_{2,3}) - 4y_1 - 6y_2 \\
\text{subject to} \quad & y_1 + y_2 = 10 \\
& 1.5x_{1,1} + 4.5x_{2,1} \le 14 \\
& x_{1,2} + 3x_{2,2} \le 14 \\
& 0.7x_{1,3} + 2.1x_{2,3} \le 14 \\
& x_{1,1}x_{1,2}, x_{1,3}, x_{2,1}, x_{2,2}, x_{2,3} \ge 0
\end{aligned}
$$

LP OPTIMUM FOUND AT STEP 3

OBJECTIVE FUNCTION VALUE

1) 30.00533

VARIABLE	VALUE	REDUCED COST
X11	9.333333	0.000000
X21	0.000000	2.500000
X12	14.000000	0.000000
X22	0.000000	1.671000
X13	20.000000	0.000000
X23	0.000000	1.166100
Y1	10.000000	0.000000
Y2	0.000000	2.000000

The new optimal solution is 30.005.

What's it worth?: *expected value of perfect information*(EVPI)

With perfect information,

* **Scenario-1**: ($x1 = 9.33$, $y1 = 10$), Profit: $30
* **Scenario-2**: ($x1 = 14$, $y1 = 10$), Profit: $30
* **Scenario-3**: ($x1 = 20$, $y1 = 10$), Profit: $30

Assuming each of these scenarios occurs with probability 1/3, his long run average profit would be

$$(1/3)(30) + (1/3)(30) + (1/3)(30) = 30$$

With his (optimal) "here-and-now" decision of ($y1 = 10$, $y2 = 0$), the long run profit is $ 30.005

This difference (30-30.005) = - 0.005 is the *expected value of perfect information*(EVPI)

***EVPI*: *Expected value of perfect information*(EVPI)** measures the value of knowing the future with certainty.

In some situations, where more information might be available through more extensive forecasting, sampling, or exploration. In these cases, *EVPI* would be useful for deciding whether to undertake additional efforts.

Example 2: The farmer's problem

Consider a farmer specializes in raising wheat, corn, and sugar beets on his 500 acres of land. During the winter, he wants to decide how much land to devote to each crop. The farmer knows that at least 200 tons (T) of wheat and 240 T of corn are needed for cattle feed. These amounts can be raised on the farm or bought from a wholesaler. Any production in excess of the feeding requirement would be sold. Over the last decade, mean selling prices have been $170 and $150 per ton of wheat and corn, respectively. The purchase prices are 40% more than this due to the wholesaler's margin and transportation costs. Another profitable crop is sugar beet, which he expects to sell at $36/T; however, there is a quota restriction on sugar beet production. Any amount in excess of the quota can be sold only at $10/T. The farmer's quota for next year is 6000 T. Based on past experience, the farmer knows that the mean yield on his land is roughly 2.5 T, 3 T, and 20 T per acre for wheat, corn, and sugar beets, respectively.

Table 1 summarizes these data and the planting costs for these crops.

	Wheat	Corn	Sugar Beets
Yield (T/acre)	2.5	3	20
Planting cost ($/acre)	150	230	260
Selling price ($/T)	170	150	36 under 6000 T
Purchase price ($/T)	238	210	10 above 6000 T
Minimum requirement (T)	200	240	- -

A farmer raises **wheat**, **corn**, and **sugar beets** on 500 acres of land. Before the planting season he wants to decide how much land to devote to each crop.

• At least 200 tons of wheat and 240 tons of corn are needed for **cattle feed**, which can be purchased from a wholesaler if not raised on the farm.

• Any grain in excess of the cattle feed requirement can be sold at $170 and $150/ton of wheat and corn, respectively.

• The wholesaler sells the grain for 40% more (namely $238 and $210 per ton, respectively.)

• Up to 6000 tons of sugar beets can be sold for $36 per ton; any additional amounts can be sold for $10/ton.

Crop yields are uncertain, depending upon weather conditions during the growing season.

Three scenarios have been identified ("good", "fair", and "bad"), each equally likely.

In this data only the yield are scenario-dependent, while in the reality the purchase prices and sales revenues from gain would be higher in year with poor yield.

To help the farmer make up his mind, we can set up the following model. Let

$x1$ = acres of land devoted to wheat,
$x2$ = acres of land devoted to corn,
$x3$ = acres of land devoted to sugar beets,

$w1$ = tons of wheat sold,
$y1$ = tons of wheat purchased,
$w2$ = tons of corn sold,
$y2$ = tons of corn purchased,
$w3$ = tons of sugar beets sold at the favorable price.
$w4$ = tons of sugar beets sold at the lower price.

Deterministic LP model is

$$\text{Min} \quad Z = 150\,x_1 + 230\,x_2 + 260\,x_3 + \left(238\,y_1 - 170\,w_1\right) + \left(210\,y_2 - 150\,w_2\right) - 36\,w_3 - 10\,w_4$$

subject to

$$\text{Land constraint: } x_1 + x_2 + x_3 \le 500$$

$$\text{Wheat constraint: } 2.5x_1 + y_1 - w_1 \ge 200$$

$$\text{Corn constraint: } 3x_2 + y_2 - w_2 \ge 240$$

$$\text{Sugar beet constraint: } w_3 + w_4 \le 20\,x_3$$

$$\text{Quota constraint: } w_3 \le 6000$$

$$x_i, y_j, w_k \ge 0$$

Deterministic solution:

After solving the above problem, the farmer obtains an optimal solution, as in Table below.

Culture	Wheat	Corn	Sugar Beets
Surface (acres)	120	80	300
Yield (T)	300	240	6000
Sales (T)	100	-	6000
Purchase (T)	-	-	-
Overall profit: $118,600			

OBJECTIVE FUNCTION VALUE

1) -118600.0

VARIABLE	VALUE	REDUCED COST
X1	120.000000	0.000000
X2	80.000000	0.000000
X3	300.000000	0.000000
Y1	0.000000	68.000000
W1	100.000000	0.000000
Y2	0.000000	41.666668
W2	0.000000	18.333334
W3	6000.000000	0.000000
W4	0.000000	16.750000

Yields: Wheat =120*2.5=300 T

Corn = 80*3= 240 T

Beet = 300* 20 = 6000 T

Does this mean that the farmer's expected revenues will actually be 118600?

A scenario representation (Stochastic approach with Discrete variables)

A first possibility is to assume some correlation among the yields of the different crops as good, fair, or bad for all crops, resulting in above average, average, or below average yields for all crops.

Let “above” and “below” average indicate a yield 20% above or below the mean yield.

The farmer wishes to know whether the optimal solution is sensitive to variations in yields.

Again, the solutions in Tables 3 and 4 seem quite natural. The optimal solution is very sensitive to changes in yields.

- The optimal surfaces devoted to wheat range from 100 acres to 183.33 acres.
- Those devoted to corn range from 25 acres to 80 acres and
- those devoted to sugar beets from 250 acres to 375 acres.
- The overall profit ranges from $59,950 to $167,667.

Long-term weather forecasts would be very helpful here. Unfortunately, as even meteorologists agree, weather conditions cannot be accurately predicted six months ahead.
The farmer must make up his mind without perfect information on yields.

Scenario	Wheat yield (Tons/Acre)	Corn yield (Tons/Acre)	Beet yield (Tons/Acre)
1. Good (Above average)	3	3.6	24
2. Fair (Average)	2.5	3	20
3. Bad (Below Average)	2	2.4	16

Scenario-1 AVERAGE	Min $Z = 150\,x_1 + 230\,x_2 + 260\,x_3 + (238\,y_1 - 170\,w_1) + (210\,y_2 - 150\,w_2) - 36\,w_3 - 10\,w_4$ subject to Land constraint: $x_1 + x_2 + x_3 \le 500$ Wheat constraint: $2.5x_1 + y_1 - w_1 \ge 200$ Corn constraint: $3x_2 + y_2 - w_2 \ge 240$ Sugar beet constraint: $w_3 + w_4 \le 20\,x_3$ Quota constraint: $w_3 \le 6000$ $x_i, y_j, w_k \ge 0$
Scenario-2 ABOVE AVERAGE	Min $Z = 150\,x_1 + 230\,x_2 + 260\,x_3 + (238\,y_1 - 170\,w_1) + (210\,y_2 - 150\,w_2) - 36\,w_3 - 10\,w_4$ subject to Land constraint: $x_1 + x_2 + x_3 \le 500$ Wheat constraint: $3x_1 + y_1 - w_1 \ge 200$ Corn constraint: $3.6x_2 + y_2 - w_2 \ge 240$ Sugar beet constraint: $w_3 + w_4 \le 24\,x_3$ Quota constraint: $w_3 \le 6000$ $x_i, y_j, w_k \ge 0$
Scenario-3 BELOW AVERAGE	Min $Z = 150\,x_1 + 230\,x_2 + 260\,x_3 + (238\,y_1 - 170\,w_1) + (210\,y_2 - 150\,w_2) - 36\,w_3 - 10\,w_4$ subject to Land constraint: $x_1 + x_2 + x_3 \le 500$ Wheat constraint: $2.x_1 + y_1 - w_1 \ge 200$ Corn constraint: $2.4x_2 + y_2 - w_2 \ge 240$ Sugar beet constraint: $w_3 + w_4 \le 16\,x_3$ Quota constraint: $w_3 \le 6000$ $x_i, y_j, w_k \ge 0$

Table 3 Optimal solution based on above average yields (+ 20%).

Culture	Wheat	Corn	Sugar Beets
Surface (acres)	183.33	66.67	250
Yield (T)	550	240	6000
Sales (T)	350	-	6000
Purchase (T)	-	-	-
Overall profit: $167,667			

Table 4 Optimal solution based on below average yields (−20%).

Culture	Wheat	Corn	Sugar Beets
Surface (acres)	100	25	375
Yield (T)	200	60	6000
Sales (T)	-	-	6000
Purchase (T)	-	180	-
Overall profit: $59,950			

The main issue here is clearly on sugar beet production. Planting large surfaces would make it certain to produce and sell the quota, but would also make it likely to sell some sugar beets at the unfavorable price. Planting small surfaces would make it likely to miss the opportunity to sell the full quota at the favorable price.

Assuming "Perfect Information", i.e., *assuming that the farmer has advance knowledge of the quality of the yield and can base his decision upon that knowledge*

Solution for scenario #1 "Good" yield

Optimal cost: 167666.6667

Stage One Variables:

i X[i]

1 183.33 ***Wheat Acres***

2 66.67 ***Corn Acres***

3 250.00 ***Beet Acres***

4 0.00

Second-stage: nonzero variables

i Y[i]

3 350.00 ***Sales of wheat***

5 6000.00 ***Sales of Beets***

Solution for scenario #2 "Fair" yield

Optimal cost: 118600

Stage One Variables:

i X[i]
1 120.00 ***Wheat Acres***
2 80.00 ***Corn Acres***
3 300.00 ***Beet Acres***
4 0.00

Second-stage: nonzero variables

i Y[i]
3 100.00 ***Sales of Wheat***
5 6000.00 ***Sales of Beets***

Solution for scenario #3 "Bad" yield

Optimal cost: 59950

Stage One Variables:

i X[i]
1 100.00 ***Wheat Acres***
2 25.00 ***Corn Acres***
3 375.00 ***Beet Acres***
4 0.00

Second-stage: nonzero variables

i Y[i]
2 180.00 ***Purchase of Corn***
5 6000.00 ***Sales of Beets***

Expected Value problem (EVP) : (General Stochastic Programming problem)

The farmer now realizes that he is unable to make a perfect decision that would be best in all circumstances.

He would, therefore, want to assess the benefits and losses of each decision in each situation.

Decisions on land assignment ($x1$, $x2$, $x3$) have to be taken now,

but sales and purchases (wi, $i = 1, . . . ,4$, $y j$, $j = 1,2$) depend on the yields.

Use a scenario index $s = 1,2,3$ corresponding to above average, average, or below average yields, respectively.

This creates a new set of variables of the form w_{is}, $i = 1,2,3,4$, $s = 1,2,3$ and yjs , $j = 1,2$, $s = 1,2,3$.

As an example, $w32$ represents the amount of sugar beets sold at the favorable price if yields are average.

Assuming the farmer wants to maximize long-run profit, it is reasonable to seek a solution that maximizes his expected profit.

This assumption means that the farmer is neutral about risk.

If the three scenarios have an equal probability of 1/3 , the farmer's problem reads as follows:

Deterministic Equivalent of SP problem directly, without decomposition

Min

$$Z = 150\,x_1 + 230\,x_2 + 260\,x_3 + \frac{1}{3}\left(238\,y_{11} - 170\,w_{11} + 210\,y_{21} - 150\,w_{21} - 36\,w_{31} - 10\,w_{41}\right)$$
$$+ \frac{1}{3}\left(238\,y_{12} - 170\,w_{12} + 210\,y_{22} - 150\,w_{22} - 36\,w_{32} - 10\,w_{42}\right)$$
$$+ \frac{1}{3}\left(238\,y_{13} - 170\,w_{13} + 210\,y_{23} - 150\,w_{23} - 36\,w_{33} - 10\,w_{43}\right)$$

This model of stochastic decision program is known as the ***extensive form*** of the stochastic program because it explicitly describes the second-stage decision variables for all scenarios.

subject to

Scenario 1:

Land constraint: $x_1 + x_2 + x_3 \le 500$

Wheat constraint: $3x_1 + y_{11} - w_{11} \ge 200$

Corn constraint: $3.6x_2 + y_{21} - w_{21} \ge 240$

Sugar beet constraint: $w_{31} + w_{41} \le 24\,x_3$

Quota constraint: $w_{31} \le 6000$

Scenario 2:

Wheat constraint: $2.5x_1 + y_{12} - w_{12} \ge 200$

Corn constraint: $3x_2 + y_{22} - w_{22} \ge 240$

Sugar beet constraint: $w_{32} + w_{42} \le 20\,x_3$

Quota constraint: $w_{32} \le 6000$

Scenario 3:

Wheat constraint: $2x_1 + y_{13} - w_{13} \ge 200$

Corn constraint: $2.4x_2 + y_{23} - w_{23} \ge 240$

Sugar beet constraint: $w_{33} + w_{43} \le 16\,x_3$

Quota constraint: $w_{33} \le 6000$

$x_i, y_j, w_k \ge 0$

LP OPTIMUM FOUND AT STEP 0

OBJECTIVE FUNCTION VALUE

1) -108390.2

VARIABLE	VALUE	REDUCED COST
X1	170.000000	0.000000
X2	80.000000	0.000000
X3	250.000000	0.000000
Y11	0.000000	22.666000
W11	310.000000	0.000000
Y21	0.000000	20.000000
W21	48.000000	0.000000
W31	6000.000000	0.000000
W41	0.000000	0.961771
Y12	0.000000	22.663002
W12	225.000000	0.000000
Y22	0.000000	17.665834
W22	0.000000	2.334166
W32	5000.000000	0.000000
W42	0.000000	8.670000
Y13	0.000000	22.663002
W13	140.000000	0.000000
Y23	48.000000	0.000000
W23	0.000000	20.000000
W33	4000.000000	0.000000
W43	0.000000	8.670000

Optimal Solution

(Found by solving deterministic equivalent problem directly, without decomposition)

```
----------------------------------------
```

Total cost: 108390

Stage One Variables:

i variable value

1 X[1] 170 ***Wheat Acres***

2 X[2] 80 ***Corn Acres***

3 X[3] 250 ***Beets Acres***

4 slack 1 0

```
----------------------------------------
```

The top line of Table 5 gives the planting areas, which must be determined before realizing the weather and crop yields. This decision is called the ***first stage***. The other lines describe the yields, sales, and purchases in the three scenarios. They are called the ***second stage***. The bottom line shows the overall expected profit.

Table 5 Optimal solution based on the stochastic model (1.2).

	Culture	Wheat	Corn	Sugar Beets
First stage	Area (Acres)	170	80	250
s =1 Above	Yield (T)	510	288	6000
	Sales (T)	310	48	6000
	Purchase (T)	-	-	-
s = 2 Average	Yield (T)	425	240	5000
	Sales (T)	225	-	5000
	Purchase (T)	-	-	-
s = 3 Below	Yield (T)	340	192	4000
	Sales (T)	140	-	4000
	Purchase (T)	-	48	-
	Overall profit: $108,390			

Second Stage

For each scenario, the optimal recourse variables are computed:

Scenario #1 "Good" yield

```
i variable value
1 Y[1]     0
2 Y[2]     0
3 W1       310  Sales of wheat
4 W2       48   Sales of corn
5 W3       6000  Sales of beets
6 W4       0
----------------------------------------
```

Scenario #2 "Fair" yield

```
i variable value
1 Y[1]      0
2 Y[2]      0
3 W1        225  Sales of wheat
4 W2        0
5 W3        5000  Sales of beets
6 W4        0
----------------------------------------
```

Scenario #3 "Bad" yield

```
i variable value
1 Y[1] 0
2 Y[2] 48  Purchase of corn
3 W1   140  Sales of wheat
4 W2   0
5 W3   4000  Sales of beets
6 W4   0
```

What’s it worth?: *expected value of perfect information*(EVPI)

With perfect information, Farmer Ted’s would plant (wheat, corn, beans).

* Good yield: (183.33, 66.67, 250), Profit: $167,667
* Average yield: (120, 80, 300), Profit: $118,600
* Bad yield: (100, 25, 375), Profit: $59,950

Assuming each of these scenarios occurs with probability 1/3, his long run average profit would be

$$(1/3)(167667) + (1/3)(118600) + (1/3)(59950) = 115406$$

With his (optimal) “here-and-now” decision of (170, 80, 250), he would make a long run profit of 108390

The difference (115406-108390) = 7016 is the *expected value of perfect information*(EVPI)

Example 3

Pop-Donuts company makes two products, donuts and cakes. It has bottlenecks in its capital used to purchase flour that is required in making both donuts and cakes, and also in labor hours used to make these two products. The goal of Pop-Donuts is to make as much profits as possible within its capacities. Currently contribution margins for a dozen donuts and a cake are $3 and $9, respectively. Pop-Donuts will adjust prices to keep same contribution margins when variable costs change. Required direct labor hours are 0.25 h per a dozen donuts and 1 h per cake. Pop-Donuts has business uncertainties and several possible scenarios in its business operations are summarized in the following:

Scenario 1: Pop-Donuts has a tight capital budget for only $15,000 of flour costs and 10,000 labor hours annually. The flour for donuts costs $0.5 per a dozen donuts and $1.2 per cake.

Scenario 2: Pop-Donuts has a modest capital budget for $17,000 of flour costs and 10,000 labor hours annually. The flour for donuts costs $0.6 per a dozen donuts and $1.3 per cake.

Scenario 3: Pop-Donuts has a very good capital budget for $20,000 of flour costs and 10,000 labor hours annually. The flour for donuts costs $0.7 per a dozen donuts and $1.4 per cake.

For Discrete Distribution:

When **discrete distributions** are used, the problem will have a finite number of scenarios. The probabilities of these three scenarios are 1/3 each. Then the problem can be formulated as follows:

Scenarios 1:	**Scenarios 2:**	**Scenarios 3:**
Max $Z_1 = 3x_1 + 9x_2$	Max $Z_1 = 3x_1 + 9x_2$	Max $Z_1 = 3x_1 + 9x_2$
subject to	subject to	subject to
$0.5x_1 + 1.2x_2 \leq 15000$	$0.6x_1 + 1.3x_2 \leq 17000$	$0.7x_1 + 1.4x_2 \leq 20000$
$2.5x_1 + x_2 \leq 10000$	$2.5x_1 + x_2 \leq 10000$	$2.5x_1 + x_2 \leq 10000$
$x_1, x_2 \geq 0$.	$x_1, x_2 \geq 0$.	$x_1, x_2 \geq 0$.

LP OPTIMUM FOUND AT STEP 0
OBJECTIVE FUNCTION VALUE

1) 101250.0

VARIABLE	VALUE	REDUCED COST
X1	15000.000000	0.000000
X2	6250.000000	0.000000

ROW	SLACK OR SURPLUS	DUAL PRICES
2)	0.000000	3.750000
3)	0.000000	4.500000
4)	15000.000000	0.000000
5)	6250.000000	0.000000

NO. ITERATIONS= 0

LP OPTIMUM FOUND AT STEP 1
OBJECTIVE FUNCTION VALUE

1) 100909.1

VARIABLE	VALUE	REDUCED COST
X1	14545.454102	0.000000
X2	6363.636230	0.000000

ROW	SLACK OR SURPLUS	DUAL PRICES
2)	0.000000	2.727273
3)	0.000000	5.454545
4)	14545.454102	0.000000
5)	6363.636230	0.000000

NO. ITERATIONS= 1

LP OPTIMUM FOUND AT STEP 0
OBJECTIVE FUNCTION VALUE

1) 102857.1

VARIABLE	VALUE	REDUCED COST
X1	17142.857422	0.000000
X2	5714.285645	0.000000

ROW	SLACK OR SURPLUS	DUAL PRICES
2)	0.000000	2.142857
3)	0.000000	6.000000
4)	17142.857422	0.000000
5)	5714.285645	0.000000

NO. ITERATIONS= 0

Solutions to these problems are shown bellow.

Scenario 1	Scenario 2	Scenario 3
Unit contribution margin Donuts: 15,000 units at $3 Cake: 6,250 units at $9 Contribution margin $101,250 Fixed expenses Manufacturing overhead $10,000 Sales and administrative $8,000 Operating income $83,250	Unit contribution margin Donuts: 14,545 units at $3 Cake: 6,366 units at $9 Contribution margin $100,902 Fixed expenses Manufacturing overhead $10,000 Sales and administrative $8,000 Operating income $82,902	Unit contribution margin Donuts: 17,142 units at $3 Cake: 5,714 units at $9 Contribution margin $102,852 Fixed expenses Manufacturing overhead $10,000 Sales and administrative $8,000 Operating income $84,852

Mean value of operating income $83,668