Stochastic Optimization
Uncertainty and Modeling Issues

Alireza Ghaffari-Hadigheh

Azarbaijan Shahid Madani University (ASMU)

hadigheha@azaruniv.edu

Fall 2017

Overview

Probability Spaces and Random Variables

- Several parameters of a problem can be considered uncertain and are thus represented as random variables.
- Uncertainty is represented in terms of random experiments with outcomes denoted by ω. The set of all outcomes is represented by Ω .
- The outcomes may be combined into subsets of Ω called events. We denote by $\mathcal{A}$ a collection of random events.
- To each event $A \in \mathcal{A}$ is associated a value $P(A)$, called a probability, such that $0 \leq P(A) \leq 1, P(\emptyset) = 0, P(\Omega) = 1$ and $P(A_1 \cup A_2) = P(A_1) + P(A_2)$ if $A_1 \cap A_2 = \emptyset$. The triplet $(\Omega, \mathcal{A}, P)$ is called a probability space that must satisfy a number of conditions.
- The elements $\omega \in \Omega$ are used to describe a few states of the world or scenarios.
- For a particular random variable ξ , we define its cumulative distribution $F_\xi(x) = P(\xi \leq x)$, or4 more precisely $F_\xi(x) = P(\omega | \xi \leq x)$.

Random variables

Discrete

- Takes a finite or countable number of different values.
- Described by its probability distribution, The list of possible values, $\xi_k, k \in\in K$, with associated probabilities,

$$f(\xi_k) = P(\xi = xi_k)$$

s. t.

$$\sum_{k \in K} f(\xi_k) = 1.$$

- Expectation $\mu = \sum_{k \in K} \xi_k f(\xi_k)$

Continuous:

- Can often be described through a density function $f(\xi)$.
- The probability of ξ being in an interval $[a, b]$ is obtained as

$$P(a \leq \xi \leq b) = \int_a^b f(\xi) d\xi = \int_a^b dF(\xi),$$

- The probability of a single value $P(\xi = a)$ is always zero for a continuous random variable.
- The distribution $F(.)$ must be such that

$$\int_{\infty}^{\infty} dF(\xi) = 1.$$

- Expectation $\mu = \int_{\infty}^{\infty} \xi dF(\xi)$.

Variance, Moments and α -quantile

- Variance $E[(\xi - \mu)^2]$
- r-th moment of ξ and is denoted $\bar{\xi}^{(r)} = E[\xi^r]$.
- A point η is called the α -quantile of ξ if and only if for $0 < \alpha < 1, \eta = \min\{x|F(x) \geq \alpha\}$.

Discrete Random Variables

Uniform: $U[1, n]$

- $P(\xi = i) = \frac{1}{n}$, $i = 1, \ldots, n, n \geq 1$
- expected value: $E[\xi] = \frac{n+1}{2}$
- Variance $Var[\xi] = \frac{n^2-1}{12}$

Binomial: $Bi(n, p)$

- $P(\xi = i) = \binom{n}{i} p^i (1-p)^{n-i}$, $i = 1, \ldots, n, 0 < p < 1$
- expected value: $E[\xi] = np$
- Variance $Var[\xi] = np(1-p)$

Poisson: $P(\lambda)$

- $P(\xi = i) = e^{-\lambda} \frac{\lambda^i}{i!}$, $\lambda > 0, i = 1, \ldots$
- expected value: $E[\xi] = \lambda$
- Variance $Var[\xi] = \lambda$

Continuous Random Variables

Uniform: $U[0, a]$

- $f(\xi) = \frac{1}{a}$, $0 < \xi < a, a > 0$
- expected value: $E[\xi] = \frac{a}{2}$
- Variance $Var[\xi] = \frac{a^2}{12}$

Exponential: $\exp(\lambda)$

- $f(\xi) = \lambda e^{-\lambda\xi}$, $\xi \geq 0, \lambda > 0$
- expected value: $E[\xi] = \frac{1}{\lambda}$
- Variance $Var[\xi] = \frac{1}{\lambda^2}$

Normal: $N(\mu, \sigma^2)$

- $f(\xi) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{(\xi-\mu)^2}{2\sigma^2}}$, $\sigma > 1$
- expected value: $E[\xi] = \mu$
- Variance $Var[\xi] = \sigma^2$

Gamma: $G(\alpha, \beta)$

- $f(\xi) = \frac{1}{\beta^2\Gamma(\alpha)} \xi^{\alpha-1} e^{-\frac{\xi}{\beta}}$

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx, \alpha > 0$$

- expected value: $E[\xi] = \alpha\beta$
- Variance $Var[\xi] = \alpha\beta^2$

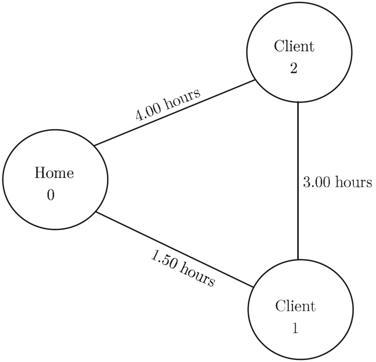
Decisions and Stages

- Stochastic linear programs are linear programs in which some problem data may be considered uncertain. Recourse programs are those in which some decisions or recourse actions can be taken after uncertainty is disclosed.
- Data uncertainty means that some of the problem data can be represented as random variables. An accurate probabilistic description of the random variables is assumed available, under the form of the probability distributions, densities or, more generally, probability measures.
- The particular values the various random variables will take are only known after the random experiment, i.e., the vector $\xi = \xi(\omega)$ is only known after the experiment.
- The set of decisions is then divided into two groups:
  - first-stage decisions: A number of decisions have to be taken before the experiment (denoted by x). The period when these decisions are taken is called the first stage.
  - second-stage decisions: A number of decisions can be taken after the experiment (denote by $y(\omega)$ or $y(\omega, x)$). The corresponding period is called the second stage.
  - The sequence of events and decisions is $x \to \xi(\omega) \to y(\omega, x)$.

An extreme example

- A traveling salesperson receives one item every day.
- Visits clients hoping to sell that item.
- Returns home when a buyer is found or when all clients are visited.
- Clients buy or do not buy in a random fashion.
- The decision is not influenced by the previous days' decisions.
- The salesperson wishes to determine the order in which to visit clients, in such a way as to be at home as early as possible.
- Time spent involves the traveling time plus some service time at each visited client.
- To make things simple, once the sequence of clients to be visited is fixed, it is not changed.
- The first stage consists of fixing the sequence and traveling to the first client.
- The second stage is of variable duration depending on the successive clients buying the item or not.

- Consider there are two clients with probability of buying 0.3 and 0.8 , respectively and traveling times (including service) as in the graph of Figure 1.
- Assume the day starts at 8 A.M.
- If the sequence is (1,2) , the first stage goes from 8 to 9:30. The second stage starts at 9:30 and finishes either at 11 A.M. if 1 buys or 4:30 P.M. otherwise.
- If the sequence is (2,1) , the first stage goes from 8 to 12:00, the second stage starts at 12:00 and finishes either at 4:00 P.M. or at 4:30 P.M.
- Thus, the first stage if sequence (2,1) is chosen may sometimes end after the second stage is finished when (1,2) is chosen if Client 1 buys the item.



Two-Stage Program with Fixed Recourse

- The classical two-stage stochastic linear program with fixed recourse (originated by Dantzig [1955] and Beale [1955]) is

$$\begin{aligned} \min \quad z = \quad & c^T x + E_\xi[\min q(\omega)^T y(\omega)] && (1) \\ s.t. \quad & Ax = b, && (2) \\ & T(\omega)x + Wy(\omega) = h(\omega) && (3) \\ & x \geq 0, y(\omega) \geq 0. && (4) \end{aligned}$$

- A distinction is made between the first stage and the second stage.
- The first-stage decisions are represented by the $n_1 \times 1$ vector x.
- Corresponding to x are the first-stage vectors and matrices c, b , and A , of sizes $n_1 \times 1, m_1 \times 1$, and $m_1 \times n_1$, respectively.
- In the second stage, a number of random events $\omega \in \Omega$ may realize. For a given realization ω , the second-stage problem data $q(\omega), h(\omega)$ and $T(\omega)$ become known, where $q(\omega)$ is $n_2 \times 1$, $h(\omega)$ is $m_2 \times 1$, and $T(\omega)$ is $m + 2 \times n = 1$.
- Each component of q, T, and h is a possible random variable.

- $N = n_2 + m_2 + (m_2 n_1)$
- $\Xi \subseteq R^N$ be the support of ξ , that is, the smallest closed subset in R^N such that $P(\Xi) = 1$.
- The dependence of y on ω is of a completely different nature from the dependence of q or other parameters on ω.
- The decisions y are typically not the same under different realizations of ω.
- They are chosen so that the constraints (3) and (4) hold almost surely (denoted a.s.), i.e., for all $\omega \in \Omega$ except perhaps for sets with zero probability.
- Difficulty:For each ω , the value $y(\omega)$ is the solution of a linear program.

Solution Steps

- For a given realization ω

$$Q(x,\xi(\omega)) = \min_y \{q(\omega)^T y | Wy = h(\omega) - T(\omega)x, y \geq 0\}. \quad (5)$$

- Define the expected second-stage value function

$$\mathcal{Q}(x) = E_\xi Q(x,\xi(\omega)). \quad (6)$$

- Solve the deterministic equivalent program (DEP)

$$\begin{aligned} \min z &= c^T x + \mathcal{Q}(x) && (7)\\ s.t. \quad & Ax = b, && (8)\\ & x \geq 0. && (9) \end{aligned}$$

- This representation of a stochastic program clearly illustrates that the major difference from a deterministic formulation is in the second-stage value function. If that function is given, then a stochastic program is just an ordinary nonlinear program.

The location problem

- Let $i = 1, \ldots, m$ index clients having demand d_i for a given commodity.
- The firm can open a facility (such as a plant or a warehouse) in potential sites $j = 1, \ldots, n$.
- Each client can be supplied from an open facility where the commodity is made available (i.e., produced or stored).
- The problem of the firm is to choose the number of facilities to open, their locations, and market areas to maximize profit or minimize costs.
- The deterministic version: so-called simple plant location or uncapacitated facility location problem.

$$\text{UFLP: } \max_{x,y} z(x,y) = -\sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{m} \sum_{j=1}^{n} q_{ij} y_{ij} \tag{4.9}$$

$$\text{s. t. } \sum_{j=1}^{n} y_{ij} \le 1\,, \qquad i = 1, \ldots, m\,, \tag{4.10}$$

$$0 \le y_{ij} \le x_j\,, \qquad i = 1, \ldots, m\,,\; j = 1, \ldots, n\,, \tag{4.11}$$

$$x_j \in \{0,1\}\,, \qquad j = 1, \ldots, n\,. \tag{4.12}$$

Parameters and Variables

- x_j: A binary variable equal to one if facility j is open and zero otherwise.
- c_j: The fixed cost for opening and operating facility j.
- v_j: The variable operating cost of facility j.
- y_{ij}: The fraction of the demand of client i served from facility j.
- t_{ij}: The unit transportation cost from j to i.
- All costs and profits should be taken in conformable units, typically on a yearly equivalent basis.
- r_i : The unit price charged to client i
- $q_{ij} = (r_i - v_j - t_{ij})d_i$: The total revenue obtained when all of client i's demand is satisfied from facility j.

In the deterministic case, it is always optimal to have each y_{ij} equal to either zero or one.)

Possible Uncertain Variables and Parameters

- Production and distribution costs may vary over time.
- Future demands for the product may be uncertain.

It is customary to consider that the location decisions x_j are first-stage decisions because it takes some time to implement decisions such as moving or building a plant or warehouse.

The firm may have full control on the distribution

- When the clients are shops owned by the firm.
- It may then choose the distribution pattern after conducting some random experiments.
- The firm may have contracts that fix which plants serve which clients,
- Or the firm may wish fixed distribution patterns in view of improved efficiency because drivers would have better knowledge of the regions traveled.

Fixed distribution pattern, fixed demand, r_i, v_j, t_{ij} stochastic

Assumptions:

- The only uncertainties: production and distribution costs and prices charged to the client.
- The distribution pattern is fixed in advance, i.e., is considered first stage.
- The second stage then just serves as a measure of the cost of distribution.
- The problem is in fact a deterministic problem in which the total revenue $q_{ij} = (r_i - v_j - t_{ij})d_i$ can be replaced by its expectation.
- Introduce extra second-stage variables w_{ij} , with the constraint $w_{ij}(\omega) = y_{ij}\ \forall\omega$.

$$\max \ -\sum_{j=1}^{n} c_j x_j + \mathrm{E}_{\xi} \sum_{i=1}^{m} \sum_{j=1}^{n} q_{ij}(\omega) w_{ij}(\omega)$$

s.t. (4.10), (4.11), (4.12), and

$$w_{ij}(\omega) = y_{ij}, \quad i = 1, \ldots, m, \qquad j = 1, \ldots, n \quad \forall\omega . \tag{4.13}$$

The second-stage objective function

By (4.13), the second-stage objective function can be replaced by

$$E_\xi \sum_{i=1}^{m} \sum_{j=1}^{n} q_{ij}(\omega) y_{ij}$$

Or

$$\sum_{i=1}^{m} \sum_{j=1}^{n} E_\xi q_{ij}(\omega) y_{ij}$$

because y_{ij} is fixed and summations and expectation can be interchanged then

$$\max \ -\sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{m} \sum_{j=1}^{n} (\mathrm{E}_{\boldsymbol{\xi}} q_{ij}(\omega)) y_{ij}$$

s.t. (4.10), (4.11), (4.12).

Fixed distribution pattern, uncertain demand

Assumptions:

- Demand is uncertain, but, for some of the reasons cited earlier, the distribution pattern is fixed in the first stage.
- The distribution costs and revenues (v_j, t_{ij}, r_i) may or may not be uncertain.
- Define y_{ij} = quantity transported from j to i.
- No longer defined as a function of the demand d_i , because demand is now stochastic.
- For simplicity, assume that a penalty q_i^+ is paid per unit of demand d_i which cannot be satisfied from all quantities transported to i and a penalty q_i^- is paid per unit on the products delivered to i in excess of d_i (the cost of inventory, for example).

Second-stage variables:

- $w_i^-(\omega)$ = amount of extra products delivered to i in state ω ;
- $w_i^+(\omega)$ = amount of unsatisfied demand to i in state ω .

$$\max - \sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{m} \sum_{j=1}^{n} (\mathrm{E}_{\boldsymbol{\xi}}(-v_j - t_{ij})) y_{ij} + \mathrm{E}_{\boldsymbol{\xi}}[- \sum_{i=1}^{m} q_i^+ w_i^+(\omega) - \sum_{i=1}^{m} q_i^- w_i^-(\omega)] + \mathrm{E}_{\boldsymbol{\xi}} \sum_{i=1}^{m} r_i d_i(\omega) \tag{4.14}$$

$$\text{s. t.} \sum_{i=1}^{m} y_{ij} \le M x_j \,, \qquad j = 1, \ldots, n \,, \tag{4.15}$$

$$w_i^+(\omega) - w_i^-(\omega) = d_i(\omega) - \sum_{j=1}^{n} y_{ij} \,, \qquad i = 1, \ldots, m \,, \tag{4.16}$$

$$x_j \in \{0,1\} \,, \; 0 \le y_{ij} \,, \; w_i^+(\omega) \ge 0 \,, w_i^-(\omega) \ge 0 \,, \quad i = 1, \ldots, m \,, \; j = 1, \ldots, n \,. \tag{4.17}$$

The objective function contains:

- The investment costs for opening plants,
- The expected production and distribution costs,
- The expected penalties for extra or insufficient demands,
- The expected revenue.

This last term is constant because it is assumed that all demands must be satisfied by either direct delivery or some other means reflected in the penalty for unmet demand. The problem only makes sense if q_i^+ is large enough, for example, larger than $E_\xi(v_j + t_{ij})$ for all$j$.

Constraints:

- Constraint (4.15) guarantees that distribution only occurs from open plants, i.e., plants such that $x_j = 1$.
- The constant M represents the maximum possible size of a plant.
- The variables y_{ij} are first-stage variables.
- In the second stage, the constraints (4.16), (4.17)
  - $w_i^+(\omega) = d_i - \sum_{j=1}^n y_{ij}$ if this quantity is non-negative . The size or capacity of plant j is $\sum_{i=1}^m d_i y_{ij}$.
  - $w_i^-(\omega) = \sum_{j=1}^n y_{ij} - d_i$ otherwise. The size or capacity of plant j is $\sum_{i=1}^m y_{ij}$.

Assumptions:

- The distribution pattern can be adjusted to the realization of the random event.
- This might be the case when uncertainty corresponds to long-term scenarios, of which only one is realized.
- Then the distribution pattern can be adapted to this particular realization.
- This also implies that the sizes of the plants cannot be defined as the sum of the quantity distributed, because those quantities depend on the random event.

model

- $x_j = 1$ if plant j is open, 0 otherwise.
- y_{ij} depend on ω with $y_{ij}(\omega)=$ fraction of demand $d_i(\omega)$ served from j.
- New variables w_j = size (capacity) of plant j , with unit investment cost g_j.

$$\max \ -\sum_{j=1}^{n} c_j x_j - \sum_{j=1}^{n} g_j w_j + \mathrm{E}_{\xi} \max \sum_{i=1}^{m} \sum_{j=1}^{n} q_{ij}(\omega) y_{ij}(\omega) \qquad (4.18)$$

$$\text{s. t.} \quad x_j \in \{0,1\}\,,\ w_j \geq 0\,, \qquad j = 1,\ldots,n\,, \qquad (4.19)$$

$$\sum_{j=1}^{n} y_{ij}(\omega) \leq 1\,, \qquad i = 1,\ldots,m\,, \qquad (4.20)$$

$$\sum_{i=1}^{m} d_i(\omega) y_{ij}(\omega) \leq w_j\,, \qquad j = 1,\ldots,n\,, \qquad (4.21)$$

$$0 \leq y_{ij}(\omega) \leq x_j\,, \qquad i = 1,\ldots,m\,,\ j = 1,\ldots,n\,, \qquad (4.22)$$

- $q_{ij}(\omega) = (r_i - v_j - t_{ij})d_i(\omega)$ now includes the deman $d_i(\omega)$.

The difference between stages and periods of times

- Consider the case of a distribution firm that makes its plans for the next 36 months.
- It may formulate a model such as (4.18)(4.22).
- The location of warehouses would be first-stage decisions, while the distribution problem would be second-stage decisions.
- The duration of the first stage would be something like six months (depending on the type of warehouse) and the second stage would run over the 30 remaining months.
- We may think of a problem over 36 periods, a two-stage model is totally relevant.
- In this case, the only moment where the number of periods is important is when the precise values of the objective coefficients are computed.

- In this example, a multistage model becomes necessary if the distribution firm foresees additional periods where it is ready to change the location of the warehouses.
- Suppose the firm decides that the opening of new warehouses can be decided after one year. A three-stage model can be constructed.
  - The first stage would consist of decisions on warehouses to be built now.
  - The second stage would consist of the distribution patterns between months 7 and 18 as well and new openings decided in month 12 .
  - The third stage would consist of distribution patterns between months 19 and 36 .

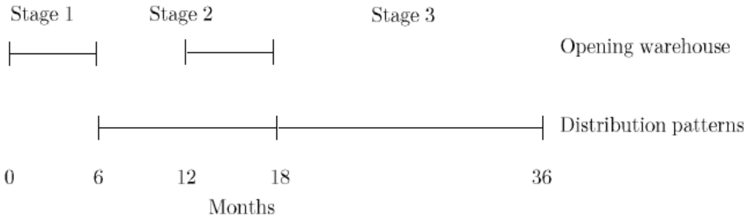


Fig. 2 Three-stage model decisions and times.

An Example of Three stage problem

- x^1 and $x^2(\omega_2)$: the binary vectors representing opening warehouses in stages 1 and 2, respectively.
- Let $y^2(\omega_2)$ and $y^3(\omega_3)$: the vectors representing the distribution decisions in Stages 2 and 3, respectively,
- ω_2 and ω_3: the states of the world in Stages 2 and 3.

$$
\begin{aligned}
\max \quad & -\sum_{j=1}^{n} c_j x_j^1 + \mathrm{E}_{\xi_2} \max\{\sum_{i=1}^{m}\sum_{j=1}^{n} q_{ij}^2(\omega_2) y_{ij}^2(\omega_2) - \sum_{j=1}^{n} c_j^2(\omega_2) x_j^2(\omega_2) \\
& + \mathrm{E}_{\xi_3|\xi_2} \max[\sum_{i=1}^{m}\sum_{j=1}^{n} q_{ij}^3(\omega_3) y_{ij}^3(\omega_3)]\} \\
\text{s. t.} \quad & \sum_{j=1}^{n} y_{ij}^2(\omega_2) \le 1 , && i = 1,\dots,m , \\
& \sum_{i=1}^{m} d_i(\omega_2) y_{ij}^2(\omega_2) \le M x_j^1 , && j = 1,\dots,n , \\
& \sum_{j=1}^{n} y_{ij}^3(\omega_3) \le 1 , && i = 1,\dots,m , \\
& \sum_{i=1}^{m} d_i(\omega_3) y_{ij}^3(\omega_3) \le M(x_j^1 + x_j^2(\omega_2)) , && j = 1,\dots,n , \\
& x_j^1 + x_j^2(\omega_2) \le 1 , && j = 1,\dots,n , \\
& x_j^1, x_j^2(\omega_2) \in \{0,1\} , && j = 1,\dots,n , \\
& y_{ij}^2(\omega_2), y_{ij}^3(\omega_3) \ge 0 , && i = 1,\dots,m , \; j = 1,\dots,n .
\end{aligned}
$$

Categorization of random events and random variables

- Place uncertainties that recur frequently on a short-term basis. (Example, uncertainty may correspond to daily or weekly demands.)
- This leads to a model, where allocation cannot be adjusted every time period.
- The expectation in the second stage somehow represents a mean over possible values of the random variables, of which many will occur.
- The expectation takes into account realizations that might not occur and many realizations that will occur.

Categorization of random events and random variables

- Place uncertainties that can be represented as scenarios, of which basically only one or a small number are realized.
- An example might be the problem of the organizers of the World Cup championship soccer game, which only occurs once every four years, to choose prices and seat allocations to maximize revenues but also to protect against possible losses.
- This consideration would also be the case in long-term models where scenarios represent the general trend or path of the variables.
- The objective function with only expected values may then be considered a poor representation of risk aversion, which is typically assumed in decision making.

Downside risk

Include risk aversion in a linear utility model under the form of a linear constraint.
The problem is to determine the type and level of production capacity at each of several locations (Eppen, Martin, and Schrage [1989]).

- Plants produce various types of cars and may be open, closed, or retooled.
- The demand for each type of car in the medium term is random.
- The decisions about the locations and configurations of plants have to be made before the actual demands are known.
- Scenarios are based on pessimistic, neutral, or optimistic realizations of demands.
- A scenario consists of a sequence of realizations for the next five years.
- Maximizes the present value of expected discounted cash flows.
- The linear constraint on risk is as follows:
    - The downside risk of a given scenario is the amount by which profit falls below some given target value.
    - It is thus zero for larger profits.
    - The expected downside risk is simply the expectation of the downside risk over all scenarios.
    - The constraint is thus that the expected downside risk must fall below some level.

An Example

$$
\begin{aligned}
\max z = c^T x + \mathrm{E}_{\xi}[\max q^T(\omega) y(\omega)] \\
\text{s. t.} \quad & Ax = b\,, \\
& T(\omega)x + Wy(\omega) = h(\omega)\,, \\
& x \geq 0\,, y(\omega) \geq 0\,. \\
& u(\xi(\omega)) \geq g - q^T(\omega)y(\omega) \\
& u(\xi(\omega)) \geq 0\,. \\
& \mathrm{E}_{\xi} u(\xi) \leq l\,,
\end{aligned}
$$

- l is some given level.
- For a problem with a discrete random vector ξ, last constraint is linear.
- This constraint is in fact a first-stage constraint as it runs over all scenarios.
- It can be used directly in the extensive form.
- It can also be used indirectly in a sequential manner, by imposing such a constraint only when needed.

A closed form expression is available for $\mathcal{D}(x)$

The stochastic queue median model (SQM)

- Locating an emergency unit.
- When a call arrives, there is a certain probability that the ambulance is already busy handling an earlier demand for ambulance service.
- In that event, the new service demand is either referred to a backup ambulance service or entered into a queue of other waiting "customers".
- The first-stage decision consists of finding a location for the ambulance.
- The second stage consists of the day-to-day response of the system to the random demands.
- Assuming a first-in, first-out decision rule, decisions in the second stage are somehow automatic.
- The quality of response, measured, e.g., by the expected service time, depends on the first-stage decision.
- When responding to a call, an ambulance typically goes to the scene and returns to the home location before responding to the next call.
- The time when it is unavailable for another call is clearly a function of the home location.

The Model

- λ : The total demand rate, $\lambda \geq 0$.
- p_i: The probability that a demand originates from demand region i , with $\sum_{i=1}^{m} p_i = 1$.
- $t(i,x)$ The travel time between location x and call i .
- On-scene service time is omitted for simplicity.
- Given facility location x , the expected response time is the sum of the mean-in-queue delay $w(x)$ and the expected travel time $\bar{t}(x)$,

$$\mathcal{Q}(x) = w(x) + \bar{t}(x)$$

$$w(x) = \begin{cases} \frac{\lambda \bar{t}^{(2)}(x)}{2(1-\lambda \bar{t}(x))} & \text{if} \quad \lambda \bar{t}(x) < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\bar{t}(x) = \sum_{i=1}^{m} p_i t(i,x), \quad \bar{t}^{(2)}(x) = \sum_{i=1}^{m} p_i t^2(i,x).$$

- The global problem

$$\min_{x \in X} \mathcal{Q}(x).$$

For a given x , $\mathcal{D}(x)$ is computable

- The deterministic traveling salesperson problem (TSP) consists of finding a Hamiltonian tour of least cost or distance.
- Each customer has a probability p_i of being present.
- A full optimization that would allow the salesperson to decide the next customer to visit at each step would be a difficult multistage stochastic program.
- A simpler two-stage model, known as a priori optimization is as follows:
  - In the first-stage, a priori Hamiltonian tour is designed.
  - In the second stage, the a priori tour is followed by skipping the absent customers.
- The problem is to find the tour with minimal expected cost (Jaillet [1988]).
- The exact representation of such a second-stage recourse problem as a mathematical program with binary decision variables might be possible in theory but would be so cumbersome that it would be of no practical value.
- On the other hand, the expected length of the tour (and thus $\mathcal{Q}(x)$) is easily computed when the tour (x) is given.

A sample Model

- c_{ij}: The distance between i and j .
- For simplicity of notation: The given tour is $\{0, 1, 2, \dots, n, 0\}$.
- $t(k)$: Expected length from k till the depot if k is present.
- We search for $Q(x) = t(0)$.
- Start with $t(n+1) = 0$ and $t(n) = c_{n0}$.
- Let $p_0 = 1$ and $c_{i,n+1} = c_{i0}$
- For $k = n-1, \dots, 0$,

$$t(k) = \sum_{r=0}^{n-k} \prod_{j=1}^{r} (1 - p_{k+j}) p_{k+r+1} (c_{k,k+r+1} + t(k+r+1))$$

- The condensed product is equal to 1 if $r = 0$.

Deterministic linear equivalent: a direct case

- An airline wishes to partition a plane of 200 seats into three categories: first, business, economy.
- The airline wishes a special guarantee for its clients enrolled in its loyalty program.
- It wants 98% probability to cover the demand of first-class seats and 95% probability to cover the demand of business class seats (by clients of the loyalty program).
- First-class passengers are covered if they get a first-class seat.
- Business class passengers are covered if they get either a business or a first-class seat (upgrade).
- Weekday demands of loyalty-program passengers are normally distributed, say ξ_F $N(16, 16)$ and ξ_B $N(30, 48)$ for first-class and business, respectively.
- The demands for first-class and business class seats are independent.

The Model

- x_1: The number of first-class seats.
- x_2 The number of business seats.
- The probabilistic constraints are

$$\begin{aligned} P(x_1 \geq \xi_F) &\geq 0.98 \\ P(x_1 + x_2 \geq \xi_F + \xi_B) &\geq 0.95. \end{aligned}$$

- Probabilistic constraints can be transformed into a deterministic linear equivalent.
- $F_F(x_1) \geq 0.98$, $F_F(.)$ denotes the cumulative distribution of ξ_F.
- $F_F(x_1) \geq 0.98$ is the same as $\frac{x_1-16}{4} \geq 2.054$ or $x_1 \geq 24.216$.
- $F_{FB}(x_1 + x_2) \geq 0.95$ is the same as $\frac{x_1+x_2-46}{8} \geq 1.645$ or $x - 1 + x - 2 \geq 59.16$.

Deterministic linear equivalent: an indirect case

The deterministic equivalent requires some transformation.

- The covering location problem.
- $j = 1, \ldots, n$: The potential locations with, $x_j = 1$ if site j is open and 0 otherwise.
- c_j: The investment cost.
- $i = 1, \ldots, m$: the clients.
- Client i is served if there exists an open site within distance t_i .
- The distance between$i$and $j = t_{ij}$.
- Define $N_i = \{j | t_{ij} < t_i\}$: The set of eligible sites for client i .
- The deterministic covering problem:

$$\min \sum_{j=1}^{n} c_j x_j \tag{7.3}$$

$$\text{s. t.} \sum_{j \in N_i} x_j \geq 1 , \qquad i = 1, \ldots, m , \tag{7.4}$$

$$x_j \in \{0, 1\} , \qquad j = 1, \ldots, n . \tag{7.5}$$

An ambulance service:

- q: The probability that no emergency unit is available at site j.
- For simplicity: This probability is the same for every site.
- The deterministic covering constraint (7.4) may be replaced by the requirement that
  P(at least one emergency unit from an open eligible site is available) $\geq \alpha$,
- α is some confidence level, typically 90 or 95%.
- The probability that none of the eligible sites has an available emergency unit is q to the power $\sum_{j \in N_i} x_j$,
- The probabilistic constraint is

$$1 - q^{\sum_{j \in N_i} x_j} \geq \alpha \,, \qquad i = 1, \ldots, m \tag{7.6}$$

$$\sum_{j \in N_i} x_j \geq b \qquad\qquad b = \left\lceil \frac{\ln(1-\alpha)}{\ln q} \right\rceil ,$$

Deterministic nonlinear equivalent: the case of random constraint coefficients

The diet problem:

- Choose among six foods (oatmeal, chicken, eggs, whole milk, cherry pie and pork with beans).
- Each food has a given serving size; for instance, a serving of eggs is two large eggs and a serving of pork with beans is 260 grams.
- Each food has therefore a known content of nutrients.
- If we take the case of protein, the content is 4 , 32 , 13 , 8 , 4 and 14 grams (grams) of proteins, respectively, for the given serving sizes.
- $x_1, \ldots, x_6$ The number of servings of each product per day:

$$4x_1 + 32x_2 + 13x_3 + 8x_4 + 4x_5 + 14x_6 \geq 55.$$

- The probabilistic constraint

$$P(4x_1 + 32x_2 + 13x_3 + 8x_4 + 4x_5 + 14x_6 \geq 55) \geq \alpha.$$

- The contents of the products are normally distributed, say $a_i\ N(\mu_i, \sigma_i^2), i = 1, \ldots, 6$.
- Assume independence between the six products.
- Then
$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6\ N(\mu, \sigma^2)$$
with
$$\mu = \mu_1x_1 + \mu_2x_2 + \mu_3x_3 + \mu_4x - 4 + \mu_5x_5 + \mu_6x_6$$
and
$$\sigma^2 = \sigma_1^2x_1^2 + \sigma_2^2x_2^2 + \sigma_3^2x_3^2 + \sigma_4^2x_4^2 + \sigma_5^2x_5^2 + \sigma_6^2x_6^2.$$
- From classical probabilistic analysis of the normal distribution implies that (7.9) is equivalent to
$$\frac{55-\mu}{\sigma} \leq z_{1-\alpha}$$
with $z_{1-\alpha}$ the $(1-\alpha)$ -quantile of the normal distribution.
- Taking $\alpha = 0.98$, the constraint reads $\frac{55-\mu}{\sigma} \leq -2.054$or $\mu \geq 55 + 2.054\sigma$.
- This constraint is non-linear and convex.

Alternative Characterizations and Robust Formulations

Relationship to Other Decision-Making Models