NTNU | Norwegian University of Science and Technology

TPK 4191 - PRODUCTION OPTIMIZATION AND CONTROL

Introduction to modelling

Jørn Vatn/August/2025

Mathemtical models

A mathematical model is a description of a system using mathematical concepts and language. The process of developing a mathematical model is termed mathematical modelling. Mathematical models are useful

- to explain a system, i.e., increase understanding regading cause and effects
- to study the effects of different *decisions*
- to make predictions about future system behaviour

Deterministic models

Equation (1) is a deterministic model for profit, $p(x)$, where it is assumed a fixed production cost of 100 Euro per unit produced, and the sale price is 300 Euro per unit but drops exponentially with an increasing sales volume. x is the production volume.

$$p(x) = 300xe^{-x/10} - 100x \tag{1}$$

There is no uncertainty in Equation (1). ▸ Excel solution

Probabilistic models

A probabilistic model is a model that enables the analyst to apply the law of total probability in an efficient way when expressing uncertainty, i.e., performing probability calculus. It is important to emphasize that a probabilistic model is not a model of the world, but it is a model used to express uncertainty regarding observables in the real world. Examples of modelling tools are Markov models and queue theory models. Equation (2) is a probabilistic model expressing the probabilityi that exactly n customers arrive a given day when the customer arrival rate per day is μ.

$$\Pr(N = n) = \frac{\mu^n}{n!} e^{-\mu} \tag{2}$$

Problem formulation - Modelling

- Structuring
- Identification of variables
  - Decision variables are those variables the decision maker can control by implementing measures, deciding upon number of items to order and so on
  - Uncontrollable variables are quantities that the decision maker can not influence
  - Uncontrollable variables are often stochastic variables
- Determination of cause and effects

Problem formulation - Modelling, cont.

- Identification of model parameters
- Specify the objective function
- Specify constraints
- Modelling
- Identification of the need of data
- Data collection and assessment of model parameters
- Run the model
- Optimization

Example - Probelm formulation

We are the project manager of an Engineer to order (ETO) project where we are approaching the final delivery of an offshore supply vessel (OSV) to an important customer

- The due date of delivery is in DD = 30 days from now on (DD = Due Date)
- If we are not able to finalize the OSV before the due date we have to pay a penalty of default equal to 10 000 Euro per day

Data has been collected for other similar ETO projects, and for this finalization phase the following data set for the number of days until finalization have been collected: 25, 35, 28, 41, 22, 31, 28 and 25

Structuring

- We have to complete the project within a given time limit
- The time it will take is random, we need to describe it
- If we exceed the due data, we have to pay penalty
- The aim is to calculate expeted penalty cost, and consider measures

Variables

- *Stochastic variable*: Let X be the number of days required to finalize the OSV. We assume that X is normally distributed with mean μ and standard deviation σ.
- We need some undestanding of probability theory........

Probabilities

- We use probabilities to express uncertainty regarding the occurence of events
- An event is something that either occurs or does not occur
- We use the notation $\Pr(\{\text{event}\})$ to express the probability that the $\{\text{event}\}$ will occur

Stochastic variables and their properties

Stochastic variables (=random quantities) are used to describe quantities which can not be predicted exactly

$$X \text{ is stochastic} \Leftrightarrow \text{ Impossible to state exactly the value of } X$$

Examples of stochastic variables are:

- X = Life time of a component (continuous)
- R = Repair time after a failure (continuous)
- T = Duration of a construction project (continuous)
- C = Total cost of a renewal project (continuous)
- N = Number of delayed trains next month (discrete)
- W = Maintenance and operational cost next year (continuous)

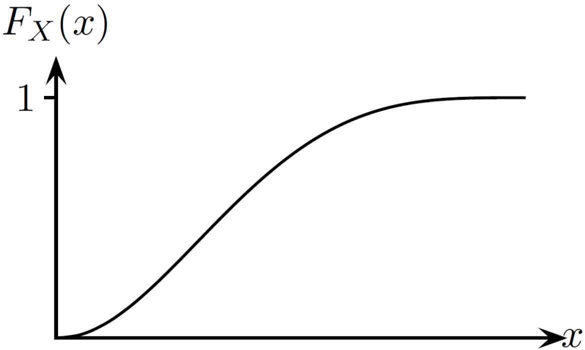
How to represent stochastic variables?

- Cumulative distribution function (CDF)
- Probability distribution function (PDF)
- Expectation
- Variance
- Standard deviation

Cumulative distribution function

A stochastic variable X is characterized by it's *cumulative distribution function* (CDF)

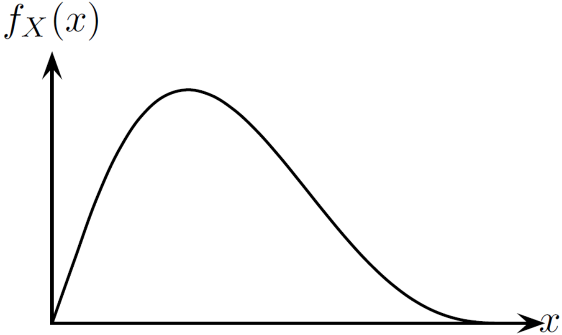
$$F_X(x) = \Pr(X \leq x)$$



Probability density function

For a continuous stochastic variable, the *probability density function* (PDF) is given by

$$f_X(x) = \frac{d}{dx} F_X(x)$$



Expectation

The expectation (mean) of X is given by

$$\mathrm{E}(X) = \begin{cases} \int\limits_{-\infty}^{\infty} x \cdot f_X(x)\, dx & \text{if } X \text{ is continuous} \\ \sum\limits_{j} x_j \cdot p(x_j) & \text{if } X \text{ is discrete} \end{cases}$$

The expectation can be interpreted as the long time run average of X, if an infinite amount of observations are available.

Variance and standard deviation

The variance of a random quantity expresses the variation in the value X will take in the long run:

$$\mathrm{Var}(X) = \begin{cases} \int\limits_{-\infty}^{\infty} [x - \mathrm{E}(X)]^2 \cdot f_X(x)\,dx & \text{if } X \text{ is continuous} \\ \sum\limits_j \left[(x_j - \mathrm{E}(X)\right]^2 \cdot p(x_j) & \text{if } X \text{ is discrete} \end{cases}$$

The standard deviation of X is given by

$$\mathrm{SD}(X) = +\sqrt{\mathrm{Var}(X)}$$

The normal distribution

X is said to be normally distributed if the probability density function of X is given by:

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where μ and σ are parameters that characterise the distribution. The mean and variance are given by:

$$\mathrm{E}(X) = \mu$$
$$\mathrm{Var}(X) = \sigma^2$$

Normal distribution, cont

The CDF cannot be found on closed form for the Normal distribution. In Excel we may use the `Normdist()` function, and in pRisk.xlsm we may use the `CDFNormal()` function.

For hand calculation it is convenient to introduce a standardised normal distribution for this purpose. We say that U is standard normally distributed if it's probability density function is given by:

$$f_U(u) = \phi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

Standardized normal distribution

We then have

$$F_U(u) = \Phi(u) = \int_{-\infty}^{u} \phi(t)\mathrm{d}t = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \mathrm{d}t$$

$\Phi(u)$ is tabulated enabeling us to find the CDF for a general normal distribution. We have that if X is normally distributed with parameters μ and σ, then $U = \frac{X-\mu}{\sigma}$ is standard normally distributed, hence

$$F_X(x) = \Pr(X \leq x) = \Pr\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = \Pr\left(U \leq \frac{X-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

Example

Let X be normally distributed with parameters $\mu = 5$ and $\sigma = 3$. We will find $\Pr(X \leq 6)$:

$$\Pr(X \leq 6) = \Pr\left(\frac{X-\mu}{\sigma} \leq \frac{6-\mu}{\sigma}\right) = \Pr\left(U \leq \frac{6-5}{3}\right)$$

$$= \Phi\left(\frac{1}{3}\right) \approx \Phi(0.33) \approx 0.629$$

Table 1: The Cumulative Standard Normal Distribution

○ + ○ = 0.33

$$\Phi(z) = \Pr(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

○ = $\Phi(0.33)$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.500	.504	.508	.512	.516	.520	.524	.528	.532	.536
0.1	.540	.544	.548	.552	.556	.560	.564	.567	.571	.575
0.2	.579	.583	.587	.591	.595	.599	.603	.606	.610	.614
0.3	.618	.622	.626	.629	.633	.637	.641	.644	.648	.652
0.4	.655	.659	.663	.666	.670	.674	.677	.681	.684	.688
0.5	.691	.695	.698	.702	.705	.709	.712	.716	.719	.722
0.6	.726	.729	.732	.732	.739	.742	.745	.749	.752	.755
0.7	.758	.761	.764	.767	.770	.773	.776	.779	.782	.785
0.8	.788	.791	.794	.797	.800	.802	.805	.808	.811	.813

Variables, cont. from the example

Overtime is considered as a means to reduce risk. If overtime is used, we assume that the extra cost is 1 000 Euro per % overtime used. Further we assume that the production rate for this last period increases linearly with the amount of overtime used.

- *Decision variable:* Let p be the percentage of overtime used, i.e., p is the decision variable.

Cause and effects

- Let μ_0 be the expected duration of this last finalization period if overtime is not used
- Causal and effect relation: We assume that the expected duration as function of p is given by $\mu(p) = \mu_0/(1 + p/100)$
- i.e., expected production rate = $r(p) = 1/\mu(p) = (1 + p/100)/\mu_0$ is linear in p

Model parameters

- The cost of overtime is $C_{\mathrm{O}}(p) = c_{\mathrm{O}}p = 1\,000p$, where $c_{\mathrm{O}} = 1\,000$ is a model parameter
- μ_0 in the equation $\mu(p) = \mu_0/(1 + p/100)$ is a model parameter
- The due date, DD is a model parameter
- The penalty per day, $C_{\mathrm{PD}} = 10\,000$ is a model parameter

Usually we use greek leters for model parameters describing stochastic variables, i.e., μ_0 and σ in the example above

Objective function

The objective function is a function we typically will maximize or minimize. The objective function is a function of the *decision variable*:

Before we can derive the objective function, we need the total expected penalty of default:

$$C_{\mathrm{PD}}(p) = \int_{x=\mathrm{DD}}^{\infty} c_{\mathrm{PD}}(x - \mathrm{DD}) f_X(x; \mu(p), \sigma^2) dx$$

Objective function

$$C_{\mathrm{PD}}(p) = \int_{x=\mathrm{DD}}^{\infty} c_{\mathrm{PD}}(x - \mathrm{DD}) f_X(x; \mu(p), \sigma^2) dx$$

where we pay $c_{\mathrm{PD}}(X - \mathrm{DD})$ if finalization date is X. Since X is random, we integrate over the probability density funcition $f_X(x|\mu(p), \sigma^2)dx$. Thus, the objective function to minimize:

$$C(p) = C_{\mathrm{O}}(p) + C_{\mathrm{PD}}(p) = c_{\mathrm{O}} p + \int_{x=\mathrm{DD}}^{\infty} c_{\mathrm{PD}}(x - \mathrm{DD}) f_X(x; \mu(p), \sigma^2) dx$$

Constraints

There could be restriction on the use of overtime, for example:

- Constraints: $p < 10$

Data and assessment of model parameters

There are unknown model parameters, here μ_0 and σ. We need to assess these. Often we use statistical methods, i.e., *parameter estimation*. Here we use:

- μ_0 =`AVERAGE()` and
- and σ =`STDEV()`

where the MS Excel functions are applied to the historical data points

Play with the model, and optimize

- What is the expected cost if we do not use overtime?
- What is the expected cost if we use 1% overtime?
- What is the optimal use of overtime?

What we need

- We need a numerical routine for $\int_{x=\mathrm{DD}}^{\infty}(x-\mathrm{DD})f_X(x;\mu(p),\sigma^2)dx$, or the best can we find a closed formula expression for the integral
- We need a numerical routine for the minimization of the cost function, or can we find a closed formula expresion for the minimum?

The backorder function - A usefull result

For the normal distribution it can be shown that the truncated expectation is given by $\int_{-\infty}^{a} xf(x;\mu,\sigma^2)dx = \mu\Phi\left(\frac{a-\mu}{\sigma}\right) - \sigma\phi\left(\frac{a-\mu}{\sigma}\right)$, where $\Phi()$ and $\phi()$ are the cumulative distribution function and probability density function for the standard normal distribution respectively. From this result it is rather easy derive the so-called backorder function:

$$B(a,\mu,\sigma) = \int_{a}^{\infty} (x-a) f_X(x;\mu,\sigma^2)\, dx$$

$$= (\mu - a)\left[1 - \Phi\left(\frac{a-\mu}{\sigma}\right)\right] + \sigma\phi\left(\frac{a-\mu}{\sigma}\right)$$

▸ Excel solution to the optimization problem

Model verification; Do we believe in our model?

We should always evaluate our mode, is it realistic, what are the assumption etc. Here:

- Is it reasonable to believe that time to finalization is normally distributed?
- Are the previous projects representative for this new one?
- Is production rate really increasing linearily with the overtime used
- Do we have data to confirm the relation between overtime and production rate?
- etc.

Thank you for your attention

NTNU | Norwegian University of Science and Technology