Descriptive Analytics
How much to produce? How much is the product?
Find the past demand data

Newsvendor problem
Example:
Time Magazine Supply chain:
- Stores were either selling out inventories (too little inventory)
- or sold only a small fraction of allocation (too much inventory)

Time magazine evaluated
- National print order
- Wholesale allotment structure
- Store distribution

What is forecasting?
Primary Function is to predict the future

Why are we interested?
Dictates the decisions we make today

Forecasts should include some distribution information
- mean and standard deviation
- range, high and low

Example usage: forecast demand for products and services
forecast inventory and capacity needs daily

What makes a good forecast
- timely and reliable
- as accurate as possible
- in meaningful units
- easy to use and comprehensive in practice

Characteristics of Forecasts
- Point forecasts are usually wrong
- Demand can be a random variable

Normal distribution is used to describe a distribution of a future relative changes in the values of stocks, FX rates
Exponential distribution can be used in characterizing time between successive arrivals of customers in service systems, e.g. call centers.

Aggregate forecasts are usually more accurate
Accuracy of forecasts erodes as we go further into the future

Subjective forecasting methods
Composites
- Sales force composites: aggregation of sales personnel estimates
- Election polling composities: sites that aggregate polls

Customer Surveys
Jury of executive opinion

Delphi method
- individual opinions are compiled and reconsidered. Repeat until overall group consensus is reached

Objective forecasting method
Causal models and time series methods

Causal Models
Let D be the demand or future outcome to be predicted and assume that there are n variables (root causes) that influence the demand
A causal model is one which demand D is formulated as a theoretical function of all those n causes
Causal models are generally intricate and complex, and need advanced tools in addition to domain expertise

Time Series methods
a collection of past values of the variable being predicted.
can be considered as a naive method. Goal is to isolate patterns in past data.
Past data might have characteristics including
trend, seasonality/cycles, randomness.

Moving averages
First calculate the mean and standard deviation

One-step forecast
Stationary data shows no trend behavior
Roughly speaking the future resembles the past

A stationary time series
$D_t = \mu + \varepsilon_t$, where μ is a constant and ε_t is a random variable with mean 0 and some deviation s

Moving average

A moving averages forecast is the arithmetic average of the n most recent observations.

We denote the moving averages method that uses n data points
MA(n)

For a one-step-ahead forecast at period t:
$F_t = (D_{t-1} + D_{t-2} + ... + D_{t-n})/n$

For moving average, a multi-step forecast is the same as the one-step forecast.

It is called moving average because the chosen data points move and are always the most recent n data points.

Predictive Statistics
Mean for Prediction = Descriptive Sample Mean
In other words, descriptive sample mean is an unbiased estimator for mean of the true demand distribution, and hence can be used for prediction.

When the data is normally distributed,
STD = $s + s/\sqrt{n}$

Advantages of moving average method
- Easy to understand
- Easy to compute
- Provides stable forecasts

Disadvantages of moving average method
Lag behind a trend
It is not a causal model, i.e. it won't explain why realizations in the future behave in a certain way.

Note that Moving Average method rops all data older than the n data points you use (Maybe recency?)
How to choose n?

You may want to give more weight to more recent data and less weight to older data

Exponential smoothing is based on this idea

Evaluation of forecast

The forecast error in period t is denoted by e_t
The difference between the forecast for demand in period t and the actual value of demand realized in t
For one step ahead forecast: $e_t = F_t - D_t$

Three ways to measure errors:
Mean absolute deviation $MAD = (1/n)\sum|e_t|$
Mean Squared Error $MAE = (1/n)\sum e_t^2$
Mean absolute Percentage error $MAPE = (1/n)\sum|e_t / D_t|x100$

Lower the errors, better the forecasting process is
Biases in Forecasts:
A bias occurs when the average value of a forecast error tends to be positive or negative

Trends and Seasonality
Data may have seasonality
Predictable annual events (cultural, weather)

Moving Averages lag trend
If there is increasing or decreasing trend in data, forecasts generated by moving averages lag behind trend.

When there is an increasing trend,
MA forecasts are usually below the demand.
When there is a decreasing trend,
MA forecasts stay above the demand.

Using regression for time series forecasting

Seasonality corresponds to a pattern in the data that repeats at regular intervals. N is the total number of seasons.

Multiplicative seasonal factors c_i where i =1 is the number of season
- $\sum_i c_i = N$
- c_i =1.25 implies 25% higher than the baseline on average
- c_i =.75 implies 25% lower than the baseline on average

In retail industry, December sales are significant, which means its sales will have a high seasonality factor in sales data.

A method of estimating seasonal factors
- Sample mean. Compute the sample mean of the entire data set

- Seasonal averages. Average the observations for the N like periods in the data.
- Seasonal factors. Divide the averages from step 2 by the sample mean

The resulting N numbers will exactly add to N and correspond to the N seasonal factors.

- De-seasonalization. To remove seasonality from a series, simply divide each observation in the data by the appropriate seasonal factor.

The resulting series will have no seasonality and is called a de-seasonalized series.

Some takeaways
Initialization issues exist for any chosen forecasting method.

There is a often a model-selection problem in how much and what data to use

Simple time-series short-term forecasting methods perform well
Long time forecasting is fraught with pitfalls since technology changes might occur.

Tracking of errors is useful for locating forecast bias.

New product problem
In the case of new products or new designs in the market, there is limited demand data.

Subjective Forecasting methods
Composites
e.g. sales force composites: aggregation of sales personnel estimates

Customer Surveys

Jury of Executive Opinion

The Dephi Method
Individual opinions are compiled and reconsidered. Repeat until overall group consensus is reached.

Empirical distribution function of forecast accuracy

A/F ratio = Actual Demand / Forecast
Start by evaluating the actual demand to forecast ratio from the N past observations.
A/F ratios measure how much the actual demands deviated from past forecasts
It helps us pin down uncertainty around current forecast.

Choosing a normal demand distribution
- Start with an initial forecast generated from subjective methods
- Evaluate the A/F ratios of the historical data

- Set the mean of the normal distribution to
Expected actual demand = Expected A/F ratio x Forecast
- Set the standard deviation of the normal distribution.
Standard deviation of actual demand = Standard deviation of A/F ratios x Forecast
Note for predictive purposes, we update the distribution to Normal with mean and standard deviation based on the moving average method.

Build an optimization model
Making best decisions in settings with low uncertainty
Converting a verbal problem description into an algebraic model: decisions, objectives, constraints
From an algebraic model to a spreadsheet implementation: optimizing with a solver
Matching demand and supply across space

Evaluating a production plan: Decision variables
Before approaching a task of finding the best production plan, or optimizing production, we must know how to evaluate any given production plan
In optimization lingo, the term decision variables describes the quantities that a decision maker can change to achieve a desired performance.

Objective function
The objective is a performance metric we want to optimize.

Constraints
In the optimization lingo, we use the term constraint to describe the requirement, e.g. the frame manufacturing hours.

Example

Maximize $150*R + 160*N$
subject to
$4*R + 5*N \leq 5610$ (frame manufacturing hours)
$1.5*R + 2.0*N \leq 2200$ (wheel and deck manufacturing hours)
$1.0*R + 0.8*N \leq 1200$ (QA and packaging hours)
R, N = integer
$R, N \geq 0$

An optimization model can have any number of decision variables and constraints but it must have one objective to be maximized or minimized.
In practice, there could be a number of quantities, key performance indicators, that a manager may want to keep track of: profit, cost, customer service levels, utilization of resources, etc.
If one of the key performance indicators, such as profit, is selected as the objective, the rest of the key performance indicators can be used in constraints.

Network Optimization

Determine the amounts of cargo to transport from each warehouse to each distribution center to minimize the total shipping cost while satisfying the warehouse supply and distribution center demand constraints.

Whether using Solvers, or commercial packages, optimization requires specifying decision variables, an objective function and constraints.

Predictive Analytics: Risk and Evaluation of Alternatives

Making Decisions in low-uncertainty vs high-uncertainty settings
In low-uncertainty settings, each particular decision produces a certain, non-random outcome, both in terms of
- the objective function value
- other key performance indicators

In a high-uncertainty environment (such as the newsvendor example) , a decision (such as the choice of a particular value for the inventory of a fashion product Q) must often be made before all the factors (such as the demand for the product D) that impact the outcome (such as profit pi) are known.

If the demand is modeled as a random variable, profit pi may also become a random variable.

In low-uncertainty settings
- For each decision, we must calculate the objective function value and determine if the decision is feasible
- Among all feasible decisions, we select one with the best objective function value

In high-uncertainty settings
- For each decision, we must know how to calculate a distribution for any key performance indicator (such as profit, resource utilization, etc.)
- When choosing the best among different decisions, we must know how to compare distributions of outcomes.

Predictive analytics provides a means to combine historical data on monthly data usage with expert judgement to come up with the probability distribution for future data usage.

Reward and risk
In dealing with uncertain outcomes it may be important to be able to calculate performance measures that can be used to compare decisions, like decisions to choose a new data plan versus staying with the old one.
When comparing decisions under uncertainty, we can then use such performance measures as an objective functions and constraints.

(Risk can be characterized by sigma in Normal distribution)

One criterion is reward: expected value of cost or profit

Making best decisions in high-uncertainty settings: steps
- Decide upon reward and risk measures
- For each competing decision, use simulation to estimate reward and risk measures
- Use reward as an objective and risk measures as constraints to find the best decision

Example
An algebraic formula: Monthly Payment for Any Value of Data Usage
P = 160 + IF(U>20, 15*(U-20), 0)
U is distributed as a normal random variable with a mean of 23, and a standard deviation of 5.

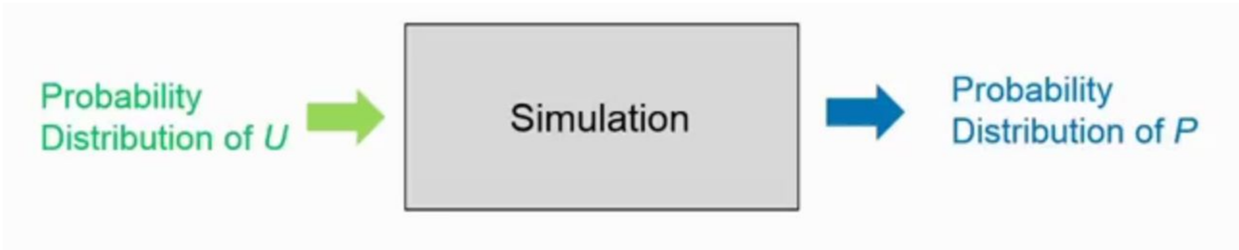
Simulation as an analytics tool
Simulation is a tool that uses a probability distribution of the "input" random variable such as data usage U to create a distribution of the "output" random variable (such as monthly payment P).

In each step of a simulation, a random instance of the input variable is generated, and the the resulting value of the output is calculated.
These simulation steps called simulation runs can be repeated as many times as necessary to generate the sample distribution of output values.

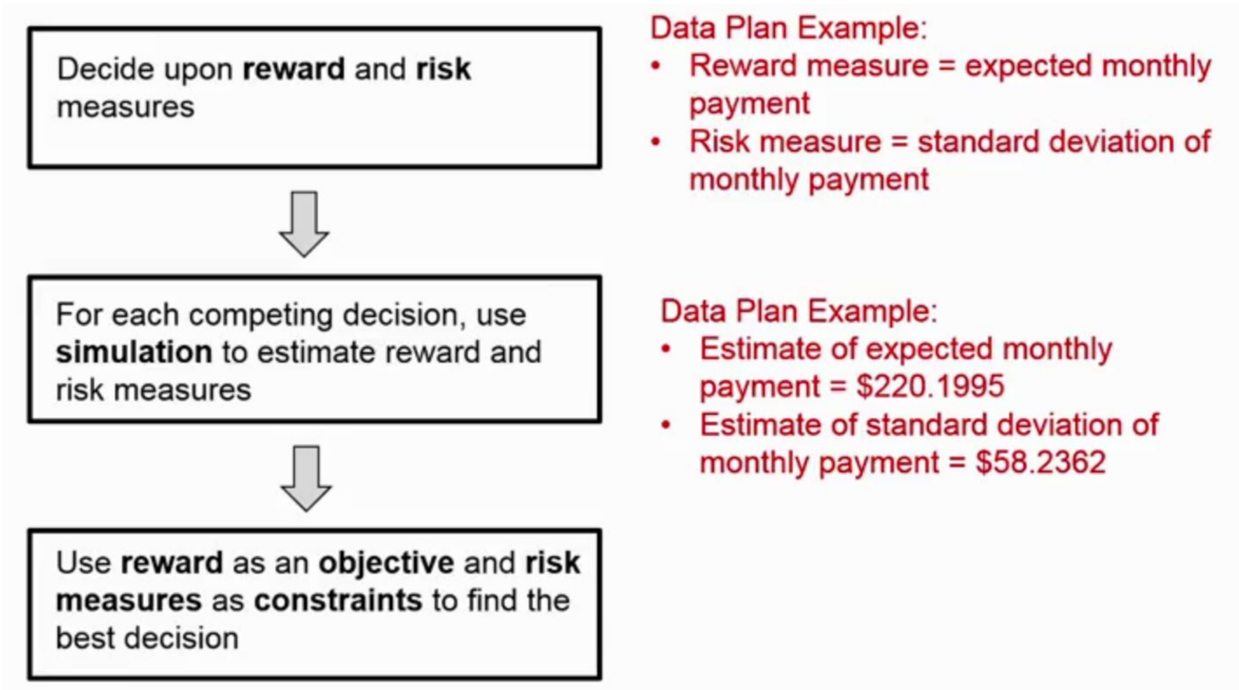
Once this sample distribution of output is generated, it can be analyzed to determine estimates for the expected value, standard deviation, etc.

Google sheets: XLMiner Analysis

Interpreting and visualizing simulation output
Histogram is used to gain intuition about the random inputs and the random outputs involved in a simulation.



Making best decisions in high-uncertainty settings: a roadmap



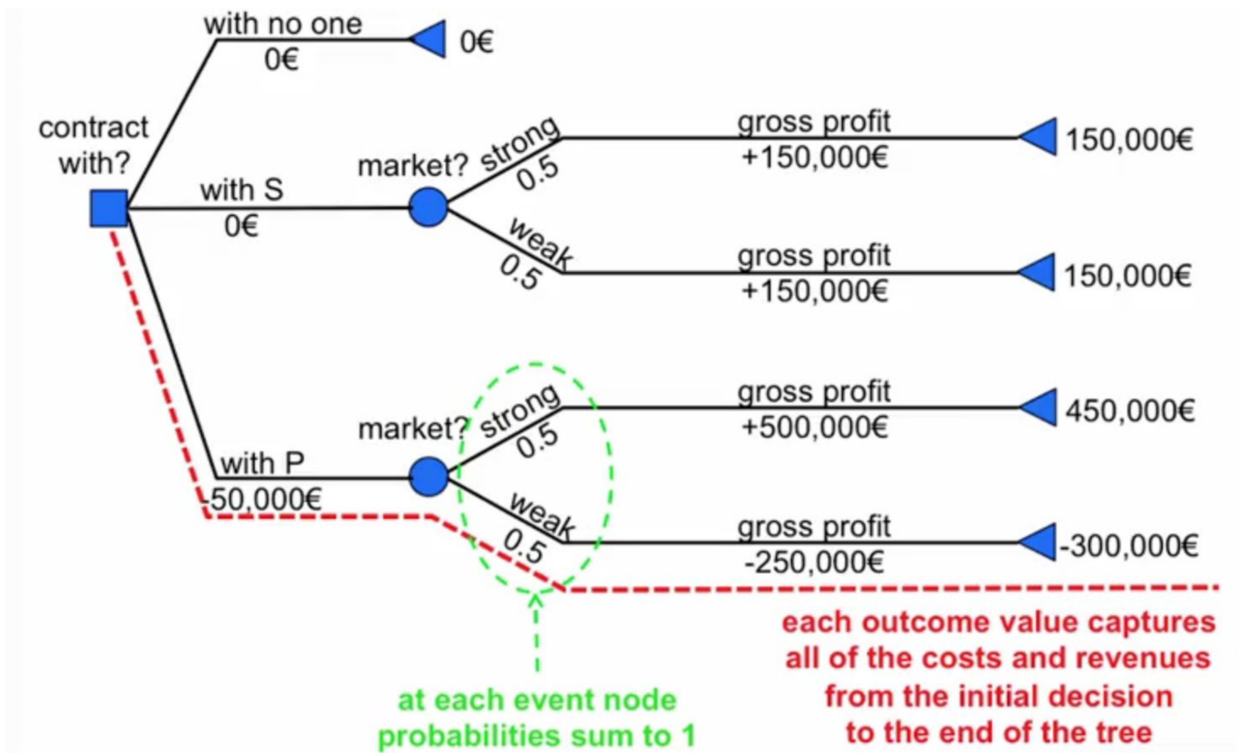
Commercial simulation packages:

http://www.orms-today.org/surveys/Simulation/Simulation.html

Prescriptive analytics: decision trees, making the best decisions in settings with high uncertainty

- "Decision" Nodes
  - Points at which a decision-maker must decide on an action
  - For IDEA this is which supplier to select, if any
- "Event" or "Chance" nodes
  - Points of uncertainty at which the outcome is random
  - For IDEA these are whether the market is weak or strong
- Outcomes
  - Payouts that occur due to specific sequences of decisions and events
  - For IDEA these are its profits
    - Profits depend on which supplier is chosen
    - And profits also depend on whether the market is weak or strong.

An example to construct a tree



Three common approaches for evaluating the options:

Maxi-min strategy

Choose the actions that maximize the minimum outcome

Avoid bad outcomes and ignores the possibility of good outcomes

Risk averse strategy

Maxi-max strategy

Choose the actiosn tha maximize the maximum outcome
Seeks good outcomes and ignores the possibility of bad outcomes
Risk seeking strategy

Maximize the expected value of the outcomes
Gives equal weight to good and bad outcomes
Risk neutral strategy

You can see the range of outcomes by looking at the tree

Using the tree to identify classic decision making strategies
Start at end, with the outcomes, and work backward to the root
At event nodes calculate the min/max/expected value
At decision nodes, cut decisions that do not maximize value

Several things to note about decision trees
- Big trees, many layers of decisions and events
- Cash flows that streams in over time should be discounted
- Where do the cash flows and probabilities come from

Past data or predictive analytics
- Sensitivity analysis to address shaky data

Find "break-even" probabilities, cash flows for decisions
- Easies to use events that have just a few discrete scenarios, but can be more complex

Using simulation with decision trees
More complex demand distribution

Using optimization together with simulation

Wrap-up for IDEA's problem
- As before, we simulated the outcomes for weak and strong markets

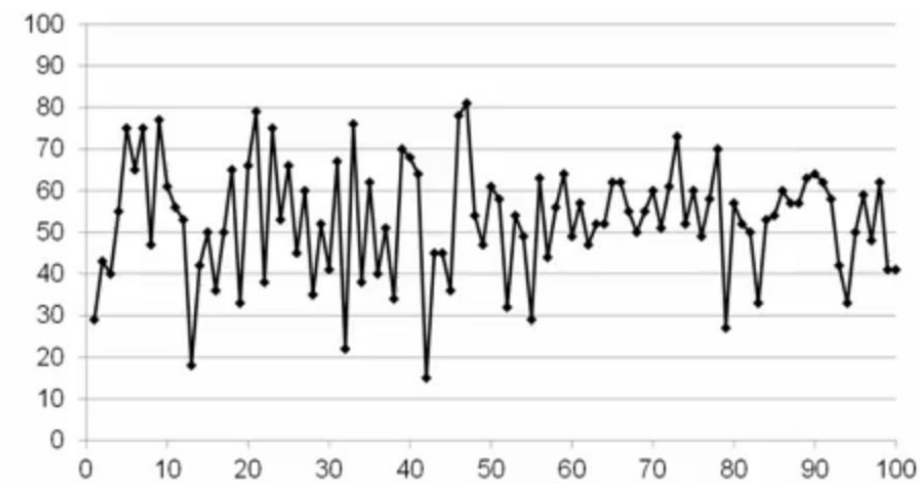
Demand model had a 50%/50% chance the market would be weak or strong
For each case we simulated uniformly distributed demand
- This time the structure of the decision problem became more complex

First IDEA needed to decide on a supplier: S,P or none
For supplier P, IDEA could then decide on an order quantity
- Rather than running separate simulation for each possible Q

We used a common set of simulated demands for all possible Qs
We optimized to find an approximately optimal Q

Back to the Newsvendor problem
Selling wodgets

Historically, demand has been variable, uncertain



How many wodgets, Q, should you order to maximize expected profit?

Solution
- Use the historical data to forecast future demand
Demand of Normal distribution
- Use the diemand forecast to drive a simulation
For a given Q simulate samples of D and calculate a pi for each sample
Calculate the average of the pi
- Use optimization to find an average-profit-maximizing Q for the sample
Objective to maximize the average profit
Deicision variable is Q
Constraints on minimum and maximum order quantity
- The optimal Q maximizes average profit for the sample
Use the optimal Q for the sample as an estimate of the optimal Q