MIS775
Decision Modelling for Business Analytics

TOPIC 9:
Quantitative Risk Analysis

AACSB ACCREDITED

Assessment task 2

Model Design Guidelines

The components listed below are examples only. You may modify the number of inputs and outputs to suit your model, provided that you meet the minimum criteria.

Hints and Minimum Requirements:

- **Fixed Inputs (minimum of 2 inputs):** Number of rooms, Housekeeping cost per room,
- **Stochastic Inputs (minimum of 4 inputs):** Daily online reservations, Late cancellations, Number of walk-ins, Daily miscellaneous expenses
- **Decision Variables (minimum of 2 decision variables, with at least 2 levels of choices for each):** Room rate/price, Overbooking rate, Late cancellation fee
- **Calculated Variables (minimum of 3 variables):** Number of occupied rooms, Daily, sales revenue, Total daily operating cost, late cancellation fee
- **Output Variables (minimum of outputs2):** Sold out status: (Yes/No), Daily profit

- **You have the flexibility in the level of complexity in your model. However, keep in mind that an overly simplistic model may lack practical value to the Sales Manager and may result in a lower assessment grade.**
- **As a Business Analyst, your goal is to design a model that meets or exceeds these requirements and delivers useful, actionable insights for the Coastal Nest Motel Sales Manager. A well-designed model will support better decision-making and contribute to improved business performance.**

Student consultation sessions – Assessment task 2

We will be holding a series of student consultation sessions to assist you with any clarifications related to your assignment. These sessions will continue **every Monday and Thursday from 7:00 PM to 7:30 PM until Monday, 19th May as given below.**

📅 Consultation Schedule:
🕖 Time: 7:00 PM – 7:30 PM (Melbourne time)
📍 Days: Every Monday and Thursday from 1st May to 19th May
🔗 **Zoom Meeting Link: provided in the Cloud Deakin Unit site**

Date	Day	Time	Week
8th May	Thursday	7 pm -7:30 pm	9
12th May	Monday	7 pm -7:30 pm	10
15th May	Thursday	7 pm -7:30 pm	10
19th May	Monday	7 pm -7:30 pm	11

Recap

- Simulation is a four step process:

  STEP 1. Establish a probability distribution for each random variable

  STEP 2. Simulate Values from the distribution

  Topic 8

  STEP 3. Replication

  STEP 4. Analyse the output data

  Topic 9

Recap

- STEP 1. Establish a probability distribution for each random variable
  - **Analyse historical data by looking at data characteristics like shape, range, discrete vs continuous, symmetric vs skewed**
  - **If unable to find a theoretical distribution that adequately fits data then use the empirical distribution derived from the data**
  - **Where there is insufficient historical data, use judgement**

Recap

- STEP 2. Simulate Values from the distribution
  - **The "what-if" approach to risk analysis involves inserting selected values for the stochastic inputs and then computing the outputs**
  - **But a far better way of evaluating risk is to use probability distributions to simulate the range of possible outcomes of a business decision, since this reflects what might be observed in practice**
  - **The data generated through simulation can then be used to quantify the risks**

Learning Objectives

- This topic focuses on risk analysis of spreadsheet-based stochastic decision models to support decision making under uncertainty
  - **Using empirical distributions to generate stochastic input data**
  - **Capturing probabilistic outcomes of multiple stochastic inputs using simulation**
  - **Simulation output analysis**
  - **Model accuracy and limitations of simulation**

Textbook reading: Chapter 10 (10.3-10.4)

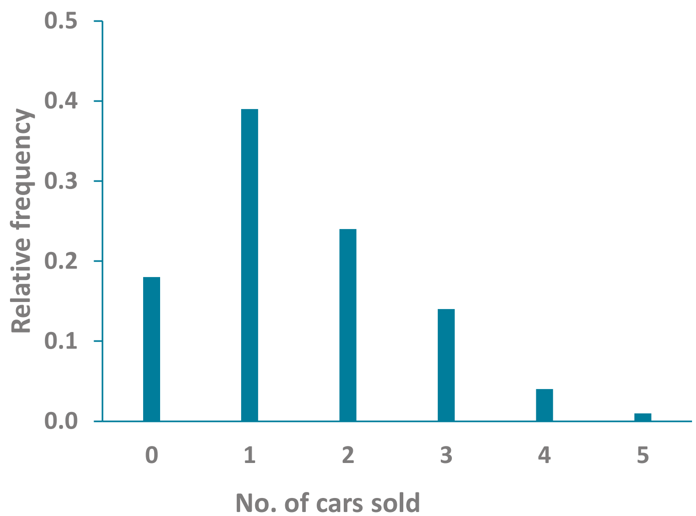
Empirical Distributions

- When historical data has no obvious fit with a theoretical distribution, we use the empirical distribution derived from the data to generate random numbers that simulate past behavior
- By using empirical distributions, input variability can be realistically simulated

Empirical Distribution – Example 1

- Using past data on daily car sales, derive the relative frequency distribution for car sales

Units Sold	No. of Days	Relative Frequency
0	54	0.18
1	117	0.39
2	72	0.24
3	42	0.14
4	12	0.04
5	3	0.01
Total	300	1.00



- There is an 18% chance that no cars will be sold during a day
- The most likely sales volume is 1, with an estimated probability of 39%
- There is a 5% chance of an outstanding sales day with four or five cars being sold

Generating Random Data from an Empirical Distribution – Example 1

- The set of random number intervals corresponding to each sales level are given in the table

- Use the following random numbers to find the units sold: 0.7128, 0.9564

  ANS: 0.7128 → 2 cars sold

  ANS: 0.9564 → 4 cars sold

Units Sold	Relative Frequency	Cumulative Relative Frequency	Interval of Random Numbers*
0	0.18	0.18	0.00-0.18
1	0.39	0.57	0.18-0.57
2	0.24	0.81	0.57-0.81
3	0.14	0.95	0.81-0.95
4	0.04	0.99	0.95-0.99
5	0.01	1.00	0.99-1.00

**From a theoretical perspective, the intervals shouldn't overlap. However, when using Excel, the chance of getting any given random number (e.g. 0.16000000) is infinitesimal so, in practice we can ignore the fact that endpoints overlap*

DEAKIN BUSINESS SCHOOL

Empirical Distribution – Example 2

- For the past 50 days, daily sales of product in a large grocery store have been recorded (to the nearest 10)

a) Determine the relative frequency for each number of units sold

b) Suppose that the following random numbers were obtained using Excel: **0.123, 0.963, 0.531, 0.809, 0.950, 0.102, 0.403, 0.458, 0.777, 0.291**

Use these random numbers to simulate 10 days of sales

Units Sold	No. of Times
30	8
40	12
50	15
60	10
70	5

Empirical Distribution – Example 2

a)

Units Sold	Relative Frequency	Cumulative Relative Frequency	Interval of Random Numbers
30	0.160	0.160	0.000-0.160
40	0.240	0.400	0.160-0.400
50	0.300	0.700	0.400-0.700
60	0.200	0.900	0.700-0.900
70	0.100	1.000	0.900-1.000

b)

Random Number	Units Sold
0.123	30
0.963	70
0.531	50
0.809	60
0.950	70
0.102	30
0.403	50
0.458	50
0.777	60
0.291	40

Empirical Distribution – Example 3

- A company sells fence posts to construction companies. They have historical sales demand for the past 26 weeks, as shown in the table
- They are interested in determining the average weekly sales using simulation for 5 weeks, so they generated random numbers for weeks 1-5 respectively: 0.63, 0.13, 0.67, 0.50, and 0.71.

Sales Category ($'000)	No. of Weeks Sold
5-10	5
11-20	7
21-30	8
31-40	6

a) Specify the random number range corresponding to each of the four historical sales levels

b) Simulate five weeks of sales and compute the resulting average weekly sales

Empirical Distribution – Example 3

a)

Sales category	Relative Frequency	Cumulative Relative Frequency	Interval of Random Numbers
5-10	5/26 = 0.19	0.19	0.00-0.19
11-20	7/26 = 0.27	0.46	0.19-0.46
21-30	8/26 = 0.31	0.77	0.46-0.77
31-40	6/26 = 0.23	1.00	0.77-1.00

b)

Week	Random Number	Sales Category	Sales Category Average
1	0.63	21-30	25.5
2	0.13	5-10	7.5
3	0.67	21-30	25.5
4	0.50	21-30	25.5
5	0.71	21-30	25.5

Average sales = (4×25.5 + 7.5)/5
= 109.5/5 = 21.9
= $21,900

STEP 3- Replication

- As the simulated input and output values change every time the spreadsheet is recalculated, we need a way of capturing and storing the simulated values
- Excel's Data Table tool is the most effective way to simulate a large number of repetitions (trials) and capture the resulting simulated output
  - **Set the column input to be any blank cell**
  - **The simulation output can be set up on a new worksheet**
  - **The resulting data table is dynamic – it will change if the sheet is recalculated, so we need to copy and paste as values**

Case Study: Replication

- Run the simulation a large number of times to generate the range of possible combinations of input and output variables

SIMULATION OF 1,000 TRIALS AT PRICE $249

Simulation trial	Direct labour cost per unit	Parts cost per unit	Demand	Contribution margin per unit ($)	First year Profit
	$46	$83.66	11788	$119.34	$406,844
1	$46	$99.18	16939	103.82	$758,578
2	$45	$95.66	9740	108.34	$55,232
3	$43	$97.19	15736	108.81	$712,250
...	...	...	...	...	...
998	$46	$81.48	17392	121.52	$1,113,532
999	$45	$93.16	16688	110.84	$849,738
1000	$45	$92.81	9757	111.19	$84,870

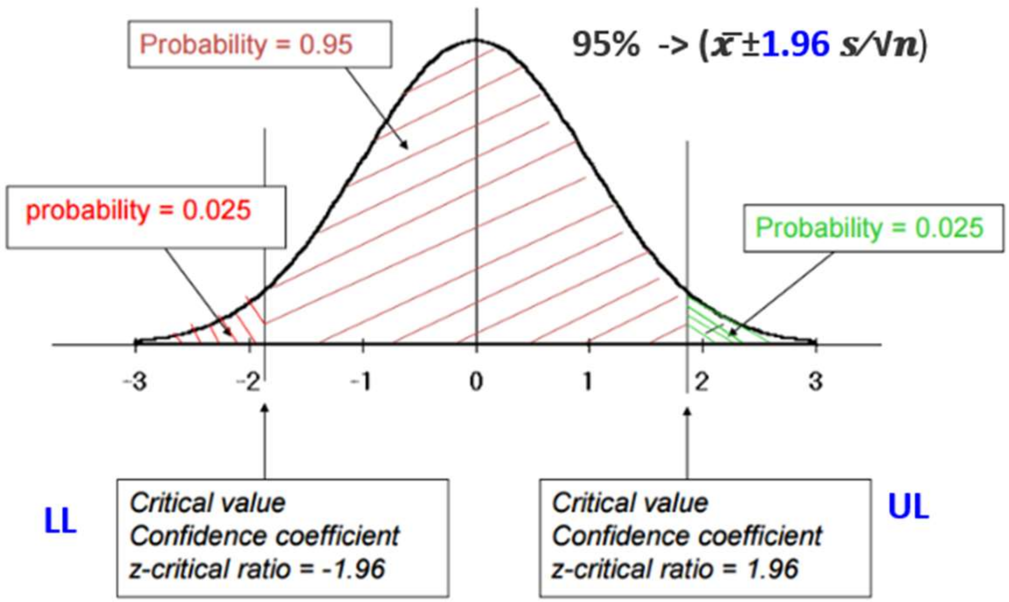
STEP 4 - Analyse the Output

- Analysis of the simulated output using descriptive statistics will give insight into the risk associated with the ‘bottom-line’:
  - **Calculate relative frequency distributions of input and output variables**
  - **Estimate the mean and standard deviation for each variable, as well as construct 95% confidence intervals for key parameters, such as the mean ($\bar{x} \pm 2\, s/\sqrt{n}$)**
  - **Identify favourable/unfavourable outcomes and their likelihoods. E.g. what is the probability of the output being negative?**
  - **Use “reverse-mapping” to identify input ranges that generated particular ranges of values of the output (associations)**

Normal distribution :
(Construct 95% confidence intervals)

Estimate the mean and standard deviation for each variable, as well as construct 95% confidence intervals for key parameters, such as the mean ($\bar{x} \pm 2\ s/\sqrt{n}$)

Confidence Coefficients for 95% Confidence Interval from standard normal distribution



Risk Profiling

- **Risk profiling:**

Using simulation output to identify particular input values or ranges of values that **give an adverse outcome** that is of interest to the decision maker

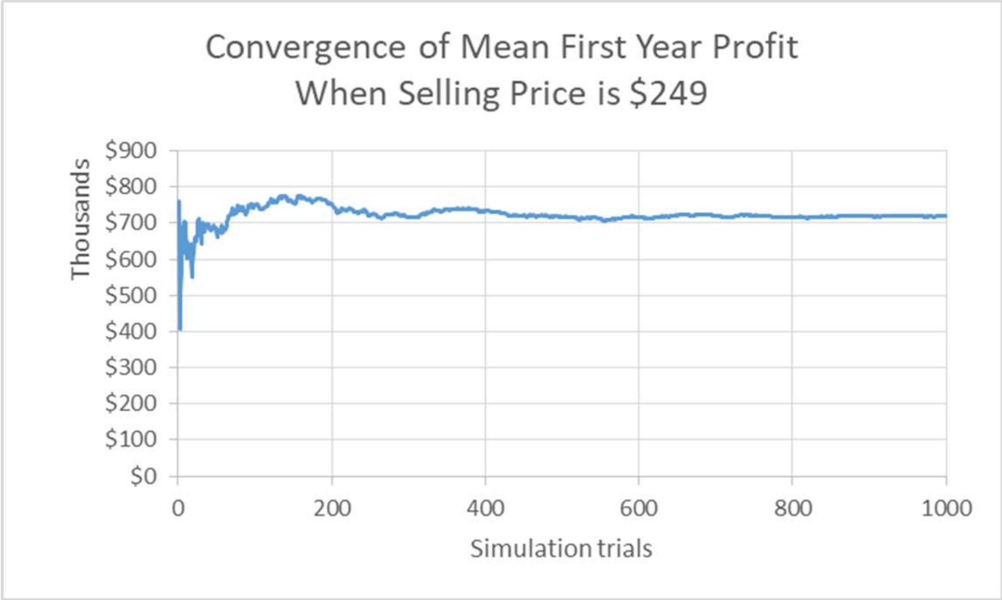
- **E.g., if a particular range of the output is unacceptable, it is possible to identify the combinations of values of the inputs that give this result so they could be avoided or reduced through an intervention, or prepared for through contingency planning**

Output Variability

- The main objective with risk analysis is to find the output variability rather than settling for an average or most likely outcome
- The variability of the 'bottom-line' can be explained through the simulated distribution
- Descriptive statistics and percentiles of the simulated distribution give insight into the risks associated with the bottom-line
- Risk management then involves either:
  - **Reducing output variability by investigating the causes of input variability and thereby attempting to reduce it**
  - **Developing contingency strategies to deal with identified risks**

Case Study: Simulation Output

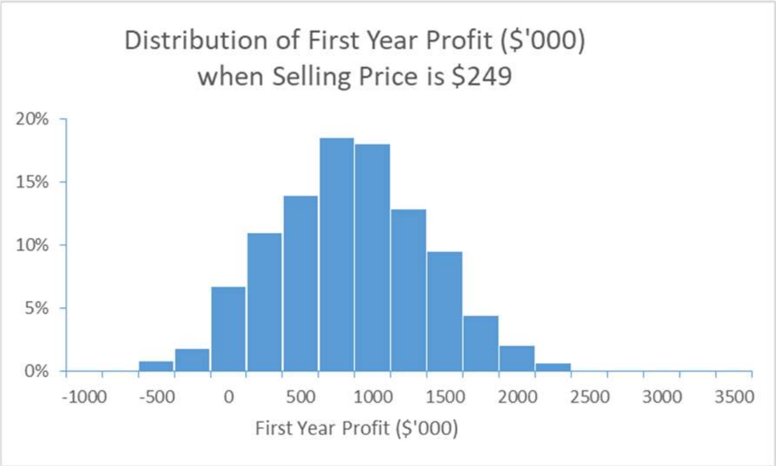
Simulation trial	First year Profit	Mean First year Profit
	$521,822	
1	$758,578	$758,578.1
2	$55,232	$406,905.2
3	$712,250	$508,686.8
4	$738,481	$566,135.5
5	$804,666	$613,841.7
…	…	…
995	$1,401,536	$719,953.0
996	$286,152	$719,517.4
997	$241,506	$719,038.0
998	$1,113,532	$719,433.3
999	$849,738	$719,563.7
1000	$84,870	$718,929.0



- This demonstrates that as the number of trials increase, the sample mean becomes a more reliable estimate of the (true population) mean of the output distribution

Case Study: Simulation Output

- Collect values from 1,000 simulation trials at each selling price, and compute summary statistics (e.g. mean, standard deviation) and include histograms
- Could also try to fit a probability distribution (e.g. Normal?) if appropriate



Case Study: Risk Profiling

- Identify input values/ranges that contribute to unfavourable/risky output values
  - **Are these input values likely to occur? If so, what are the contingency plans?**

First Year Profit intervals when Selling Price is $249

	-$1M to -$0.5M	-$0.5M to $0	$0 to $1M	$1M to $2M	Over $2M	Total
Probability	0.8%	8.5%	61.3%	28.7%	0.7%	100.0%
Max demand	4,698	9,534	19,113	27,169		

Model Errors

- **There are two main sources of errors:**
  - **Model logic errors :** the way your simulation is built might be incorrect. To avoid this, you should **test and validate** your model thoroughly **before** adding random (stochastic) inputs. Make sure the model behaves as expected under known conditions. (Week 7)
  - **Poor Input Data Quality**
    Even if your model is perfect, bad data will still give you bad results.
    - Irrelevant or incorrect data ("garbage in, garbage out")
    - Missing information
    - Outliers (unusual data points that can distort results)

Build the model carefully and use good data to get meaningful simulation results.

Model Errors

- Empirical data is sampled data – it is not meant to be complete or completely representative
- Simulation is not meant to be perfect or reproduce the reality or provide the optimal result. It helps you **understand and explore uncertain situations**. It's a tool to give you **insight**, not a **crystal ball** that predict the future with 100% accuracy.
- **Must validate your simulation output! Do a simple common sense test and ask a subject matter / domain expert!**

Where to Now?

- Our case study only considered two possible selling prices *(one decision variable with 2 levels of choices $249 and $299)*. Perhaps additional research could be undertaken to estimate the price-demand response curve. By incorporating this into the model, management could explore profit at other price points
- Demand would also depend on the levels of administrative and advertising spend. Perhaps Sanotronics management could acquire a loan so these could be recategorised as decision variables
- Can we model that – and then discover what might be the best selling price?

Summary

- Stochastic modelling produces more realistic models that capture and highlight risk
- We have discussed:
  - **Using empirical distributions to generate stochastic input data**
  - **Capturing probabilistic outcomes of multiple stochastic inputs using simulation**
  - **Conducting simulation output analysis**
  - **Model accuracy and limitations of simulation**

Next Class

- **Topic 10: Applications of Simulation Modelling – Queueing Models**
  (Ch. 12 of textbook – sections 12.1-12.5)