Defence Research and Development Canada
Recherche et développement pour la défense Canada

Monte Carlo Methods for Cost Estimation

Tania Yazbeck
DRDC CORA, Defence Economics Team
ICEAA Canada Workshop, 2018

DRDC | RDDC

Canada

Outline

- Deterministic vs simulated - The Monte Carlo principle
- Monte Carlo in cost estimating
  - Motivation
  - Best practices and common mistakes
- Implementation considerations
- Other applications of MC methods
- Conclusion
- Questions?

Deterministic vs Stochastic
The Monte Carlo principle

Definition of Monte Carlo Method

- Any method which solves a problem by generating suitable random numbers and observing that fraction of the numbers obeying some property or properties.

Weisstein, Eric W., "Monte Carlo Method". From MathWorld – A Wolfram Web Resource. http://mathworld.wolfram.com/MonteCarloMethod.html

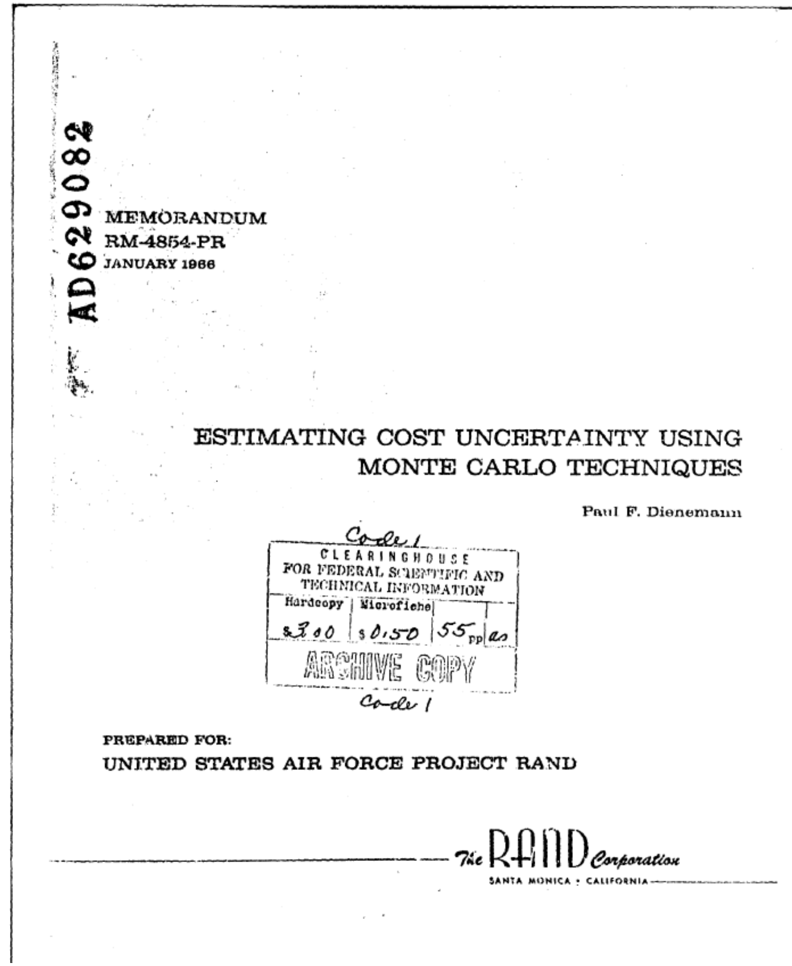
History

Monte Carlo methods are not new…

- *Metropolis, N. and Ulam, S. "The Monte Carlo Method." J. Amer. Stat. Assoc.* ***44****, 335-341, 1949.*
- However, method was used before.

Monte Carlo methods in cost estimation also not new…

- *Paul F. Dienemann, "Estimating Cost Uncertainty Using Monte Carlo Techniques", RAND Corporation, 1966.*


AD629082

MEMORANDUM
RM-4854-PR
JANUARY 1966

ESTIMATING COST UNCERTAINTY USING
MONTE CARLO TECHNIQUES

Paul F. Dienemann

Code 1

CLEARINGHOUSE
FOR FEDERAL SCIENTIFIC AND
TECHNICAL INFORMATION

Hardcopy	Microfiche		
$3.00	$0.50	55 pp	as

ARCHIVE COPY

Code 1

PREPARED FOR:
UNITED STATES AIR FORCE PROJECT RAND

The RAND Corporation
SANTA MONICA • CALIFORNIA

Illustration of the Monte Carlo principle

PROBLEM : **When rolling one die, what is the probability of getting a 4?**

- Two strategies for finding the answer:

  SOLUTION 1 - Calculate:

  Pr{X=4} = 1/# possible outcomes = 1/6

  SOLUTION 2 – Simulate:

  Roll the die many times and count how many times the outcome is 4

  $\widehat{Pr}$ {X=4} = (#observations where X=4)/(total number of trials)

- By the law of large numbers, simulated solution will approach calculated solution as the number of trials increases.

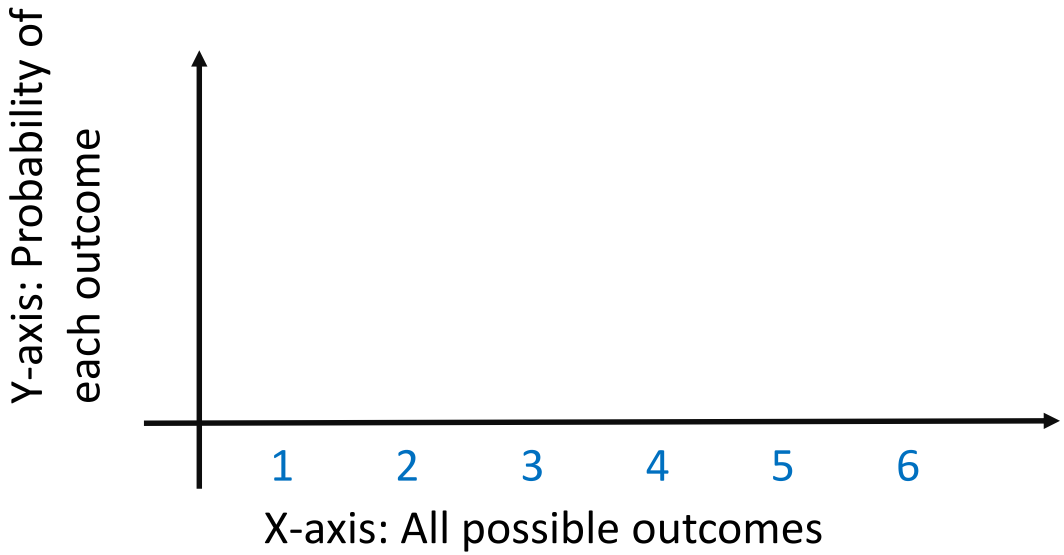
Simple project planning example

	Costs
Labour	10,000
Materials	15,000
Permits	1,000
Total	26,000
	Time
Design	1 week
Phase 1	3 weeks
Phase 2	5 weeks
Finishing	2 weeks
Total	11 weeks

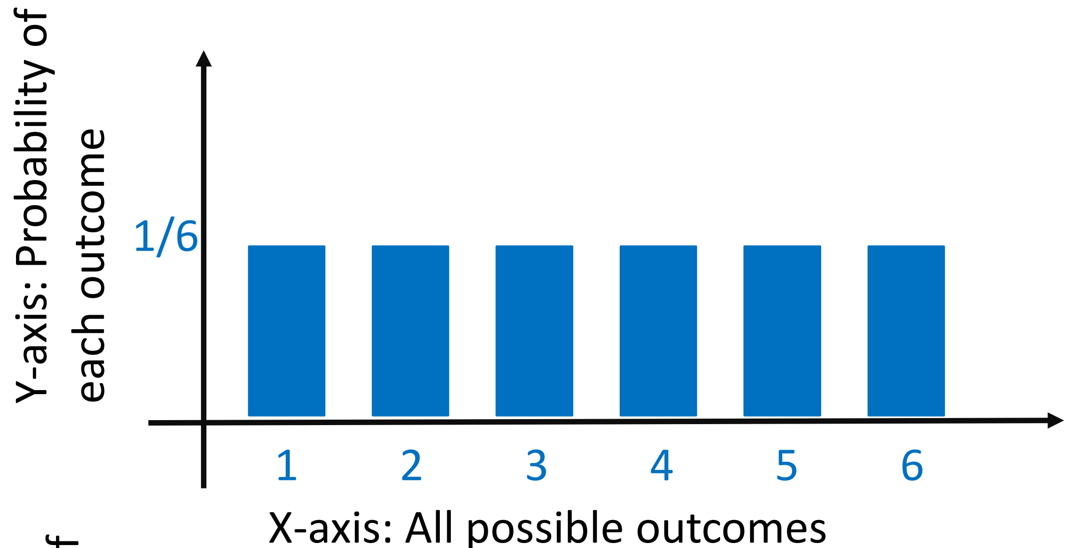
Monte Carlo Algorithm

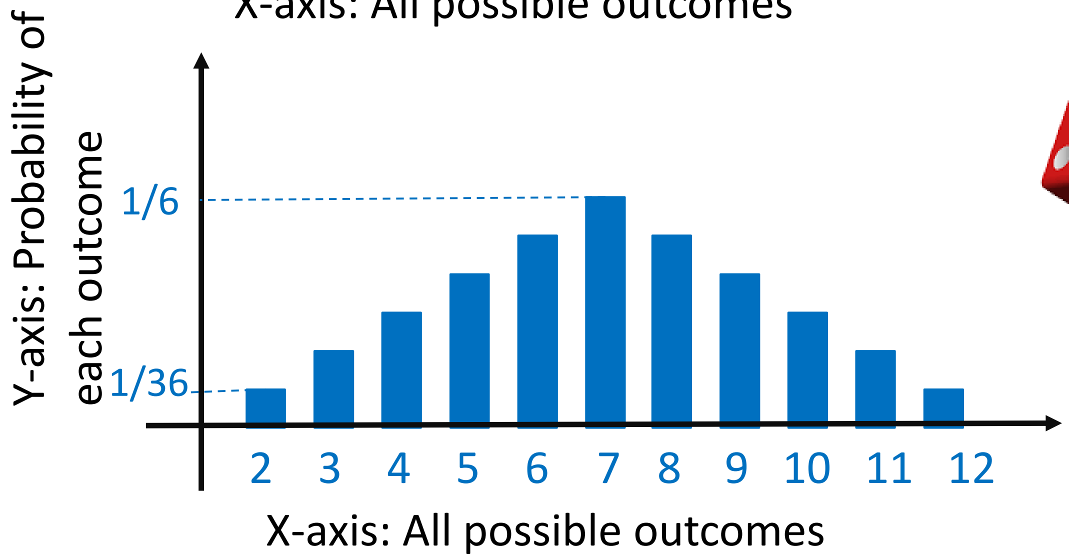
1. Identify model variables that are subject to uncertainty
2. Identify for each variable the probability distribution that best represents the type of uncertainty that it is subject to
3. Evaluate model using values randomly sampled from chosen probability distributions
4. Repeat many times
5. Aggregate results

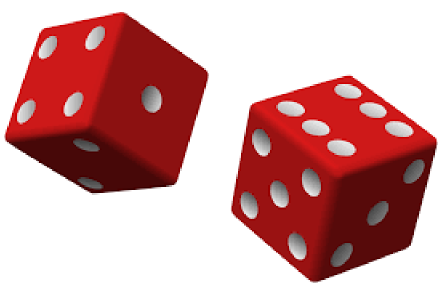
“Sampling from a probability distribution”



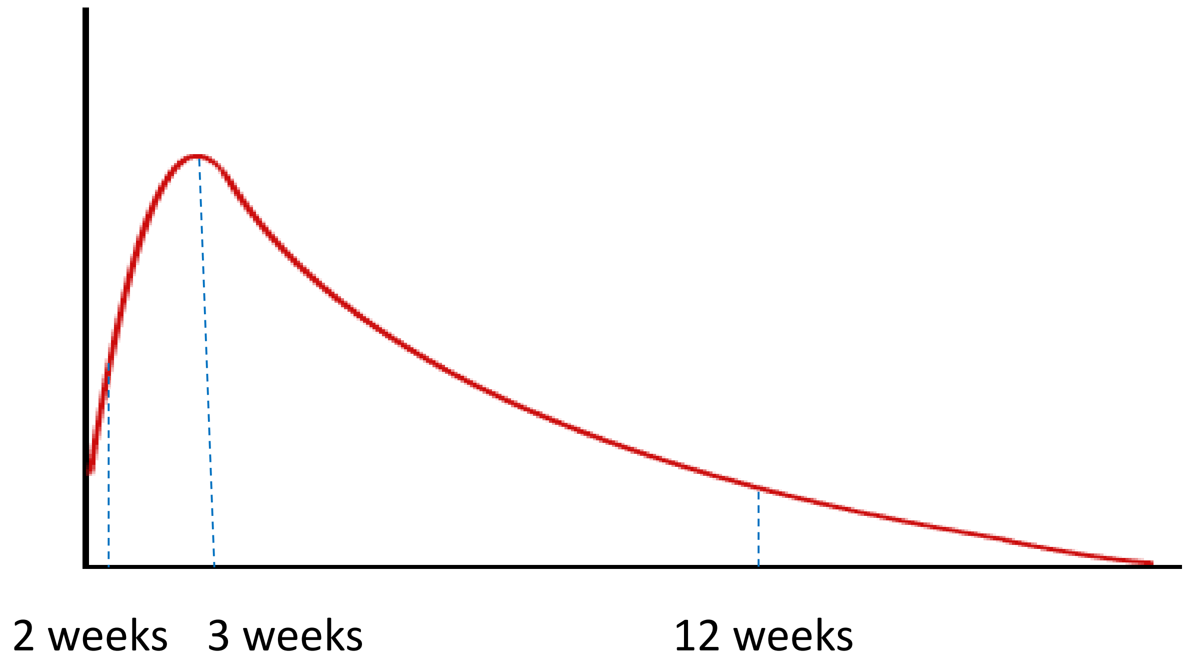
“Sampling from a probability distribution”







“Sampling from a probability distribution”



Duration of ‘phase 2’ in earlier example

Monte Carlo method in Cost Estimating

Motivation

- MC methods provide a systematic way of examining the risk and uncertainty associated with each cost element.
- Account for various forms of uncertainty.
- Can combine top-down and bottom-up estimation methods.
- Provide a statistically sound method for aggregating the cost uncertainty into a project-level (system-level) estimate.
- Result in a final project estimate that is represented as a range of possible costs with associated confidence intervals.
- Dynamic model that can be updated as new information becomes available or environment changes (for example).

Sources of cost risk and uncertainty

- Numerous sources of cost risk and uncertainty such as:
  - Technical requirements
  - Technological changes
  - Environmental variables

- Probability distributions can be used to model data that is:
  - Subject to external variables (economy, weather, labour rates, …)
  - Subject to internal variables (HR, IT considerations…)
  - Unreliable
  - Missing
  - Ambiguous

4 Best Practices and 1 Common Mistake

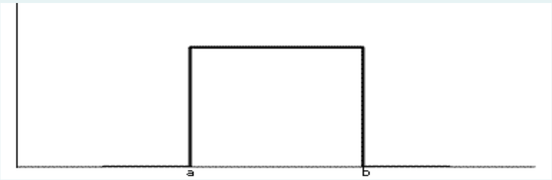
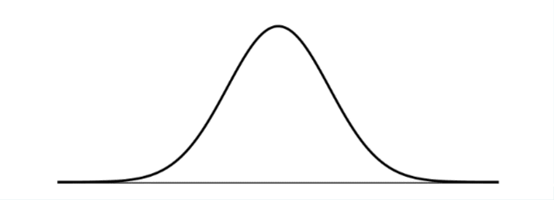
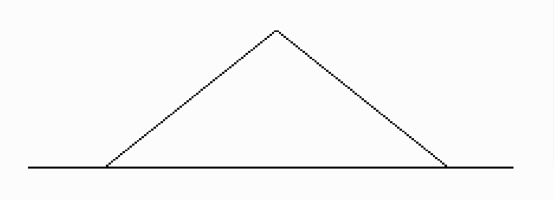
Best practices - 1

- Calculate whenever possible

- Because uncertainty of one input is compounded by the uncertainty of all other inputs, simulating already known quantities unnecessarily enlarges the prediction interval of the final cost/schedule.

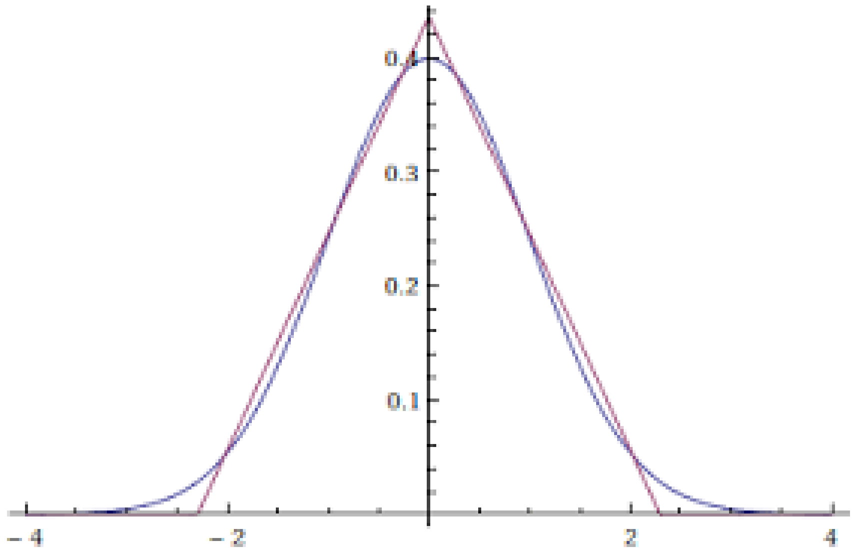
Best practices - 2

- <u>Find the appropriate probability distribution</u>
- When in doubt:

Distribution	
Uniform	
Normal	
Triangular	

Best practices - 3

- <u>Be careful not overuse the triangle distribution.</u>
- Useful when little is know of a cost element.
- Ideal for eliciting expert judgement.
- But… Ignores most extreme outcomes because bounded.
- Not a good way to summarize existing data.



Src: Math.stackexchange.com

Best practices - 4

- Repeat simulation enough times
- How many is enough?
- Depends on:
  - Number of simulated variables
  - Shapes of distributions
  - Desired accuracy

- Rule of thumb:

  At least 1,000 runs

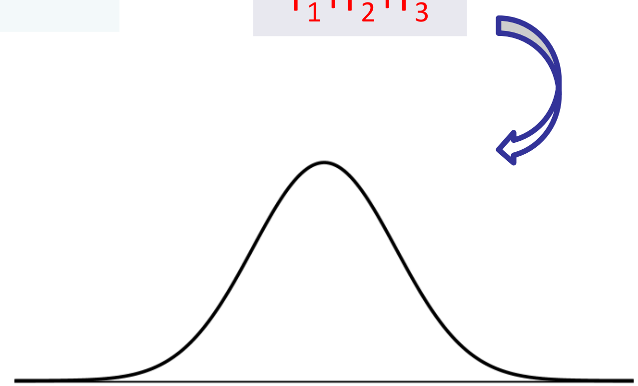
  Preferably 10,000
- Ideally, go for convergence

A common mistake…

- Cannot add min, max or most likely values of probability distributions.

	Min	Most likely	Max
Labour	8,000	10,000	12,000
Materials	14,000	15,000	25,000
Permits	1,000	1,000	1,250
Total	---	---	---

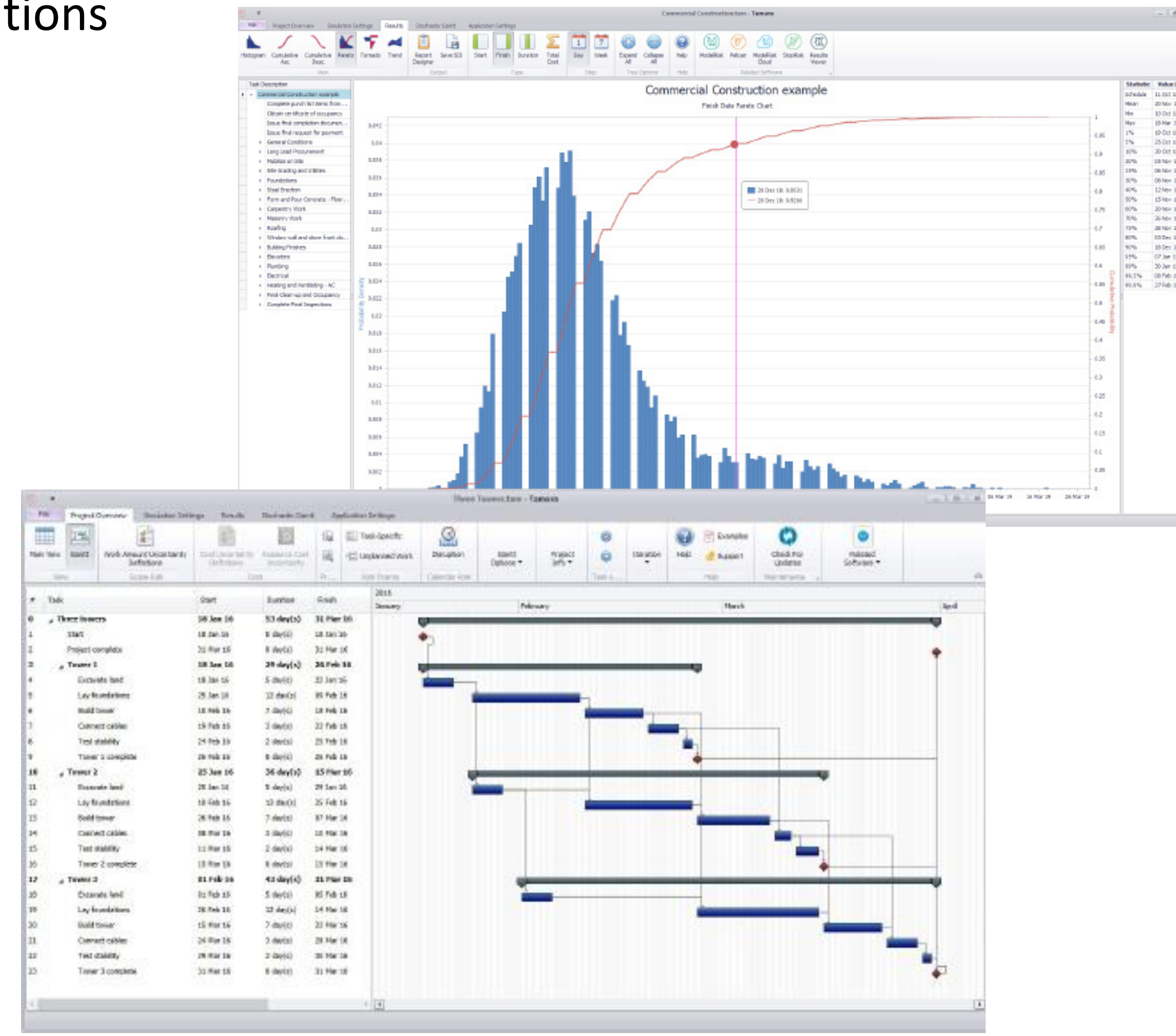
MC run i
r_1
r_2
r_3
$r_1+r_2+r_3$



Implementation considerations

Software solutions

- A number of software solutions
- Examples:
  - @Risk
  - CrystalBall
  - ModelRisk
- Excel add-ons



Programming solutions

- Examples:
  - MATLAB
  - R Monte Carlo package
  - Python

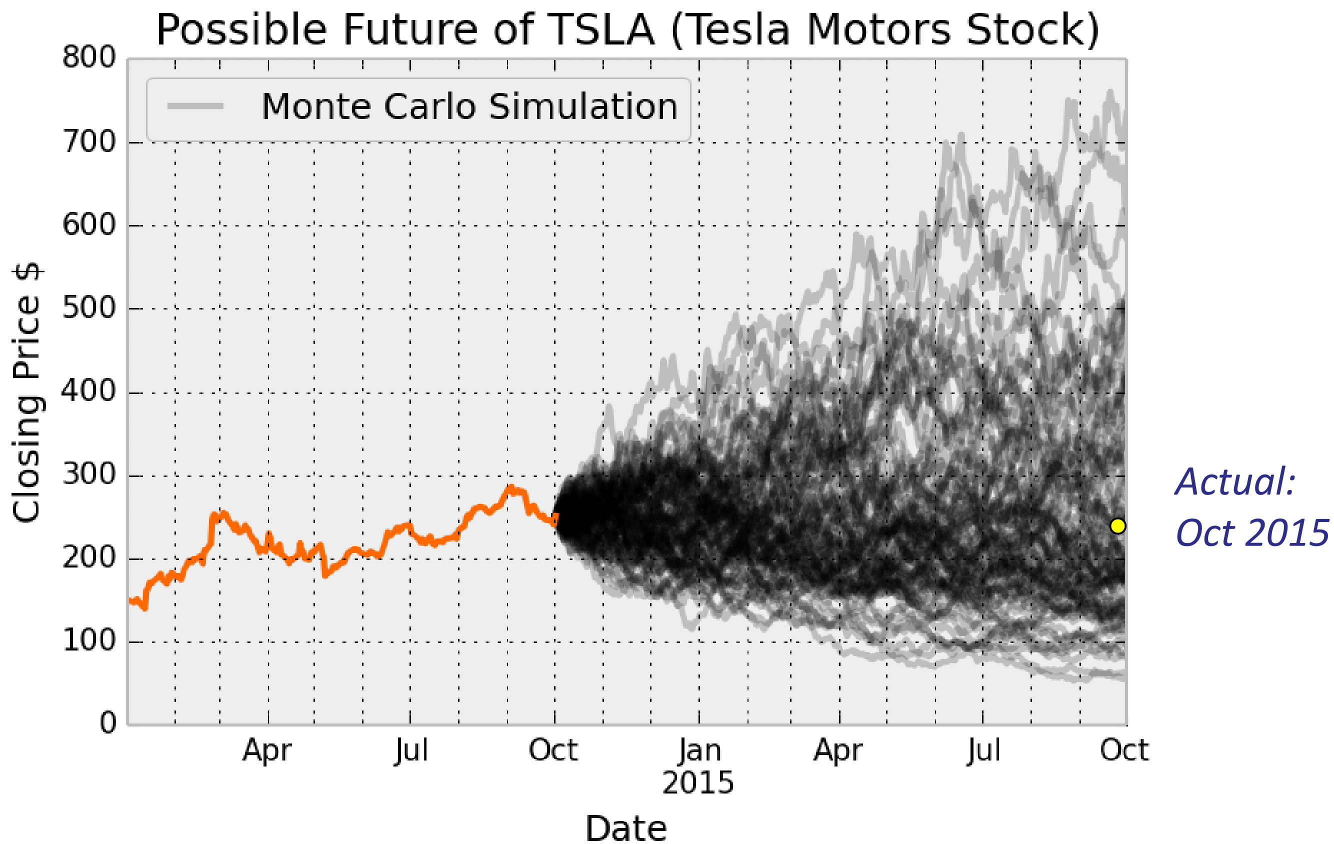
```
library(MonteCarlo)
# define parameters
…
# run simulation:

MC_result<-MonteCarlo(func=ttest, nrep=1000,
                      param_list=param_list)
```

Monte Carlo simulation in R – example commands

Other applications of Monte Carlo method

Finance

- Modeling of expected stock returns
- Portfolio management (VaR)
- Pricing of derivatives



Src: https://github.com/balzer82/Pandas-Stoxx-Europe-600

Artificial Intelligence

- Everywhere reasoning under uncertainty is required…
- Machine learning algorithms
- Robotics (search location in space)
- Games

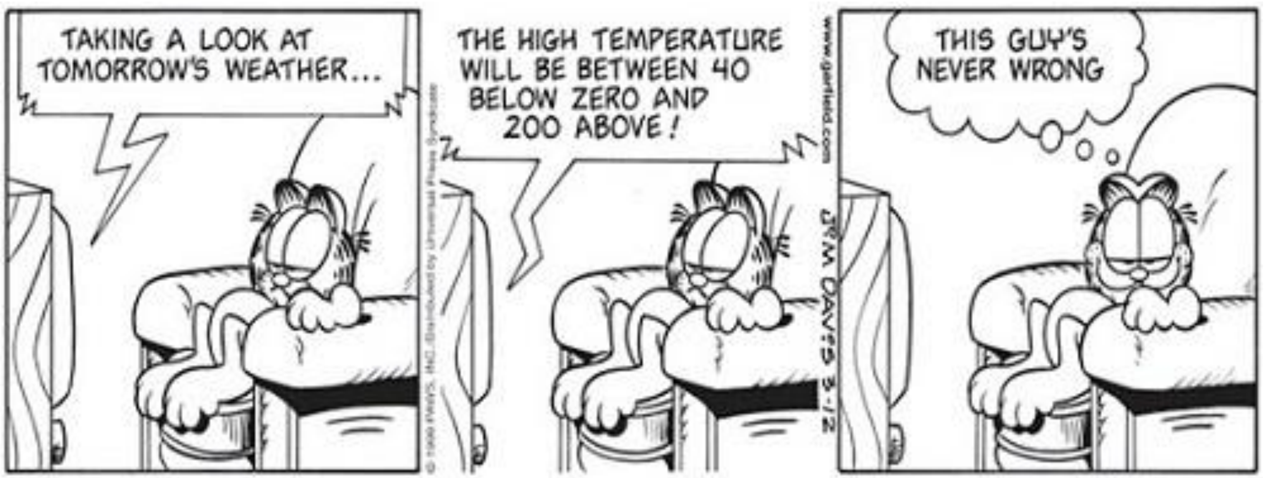
Src: http://aigamedev.com/open/editorial/2013-awards/

Conclusions

Conclusion

- Monte Carlo methods provide a way to systematically describe and model the various risks and uncertainties affecting a project.
- Lead to more reliable cost estimates by producing a range of likely outcomes for the final cost or schedule of a project rather than a point estimate.
- Lead to a better understanding of the drivers of cost and hence can help in setting priorities or defining mitigation strategies.
- The method is simple, intuitive and easily implemented,
- However, care must be taken when selecting probability distributions.
- Inappropriate choice of sampling distributions can lead to inaccurate or overly large range of probable outcomes.

Questions



https://www.kaggle.com/timolee/finding-trees-through-the-forest-a-rf-dissection