Statistical Inference for Engineers

Maranello, 19th March 2012

Roger Frigola

My Background

- Degrees in General and Aerospace Engineering: dynamics, control & aero
- AIRBUS (1.5 years) Fluid-Control-Structure Interaction
- McLaren F1 Team (4 years) Simulation Algorithms, Modelling, Optimisation, Design of Wind Tunnel Experiments.
- Now working on a Statistical Machine Learning PhD at Cambridge.

UNCERTAINTY

Some Sources of Uncertainty

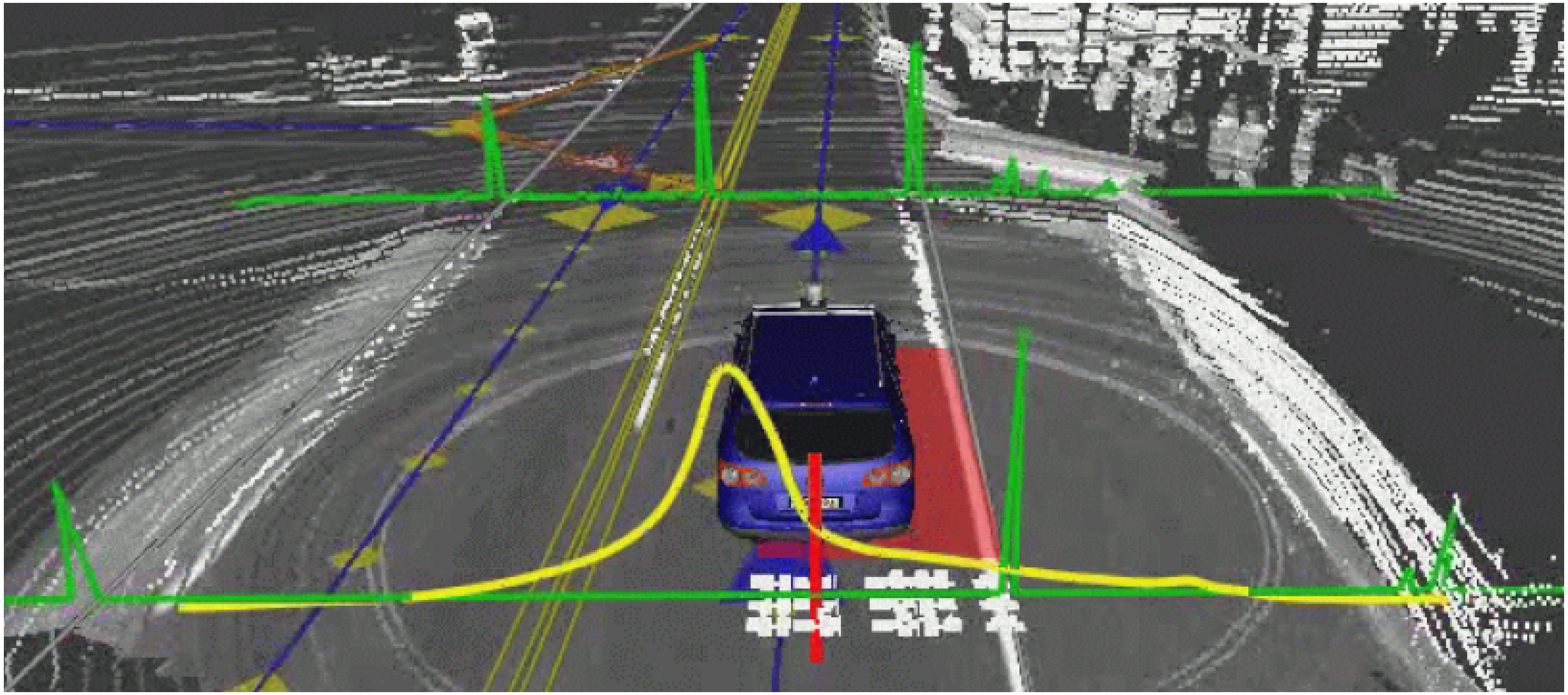
- Uncertainty due to noisy or faulty sensors.
- Uncertainty due to a parameter not being directly observed.
- Uncertainty due to test rigs of unknown fidelity.
- Uncertainty due to unmodelled effects.
- Uncertainty due to a limited amount of time to run numerical simulations.
- and many more...

Inference, Statistics and Machine Learning
Who is who?

- Inference: task of finding a coherent interpretation of incoming observations that is consistent with both the *observations* and the *prior information* at hand.

- Statistics: related to Mathematics. Focuses on proofs and guarantees.

- Machine Learning: related to Computer Science. Focuses on practical implementations.



Minimising the Squared Errors

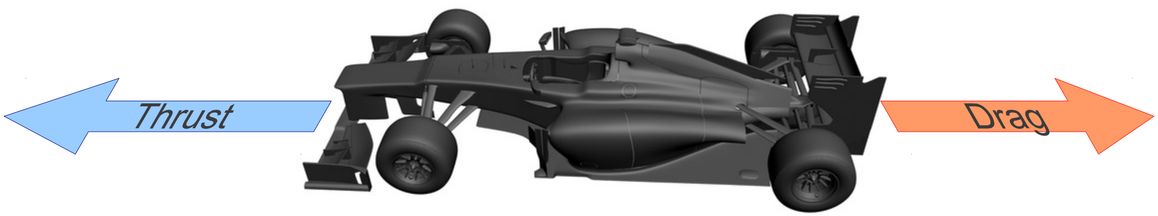
- The engineers' preferred inference tool.

- Example: infer power and drag given 100 noisy accelerometer measurements.
  - For instance using a linear model:

$$a = \left(\frac{1}{mV}\right) P - \left(\frac{1}{2m} \rho V^2 S\right) Cd$$

  - Minimise the squared errors:

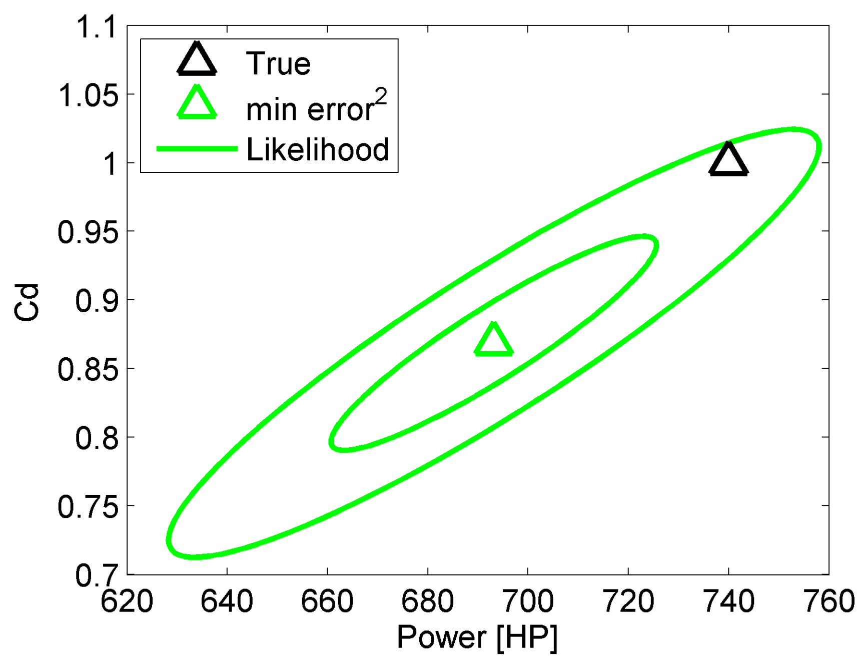
$$\sum \left(aMeasured_i - aModel_i(P, Cd)\right)^2$$



Minimising the Squared Errors

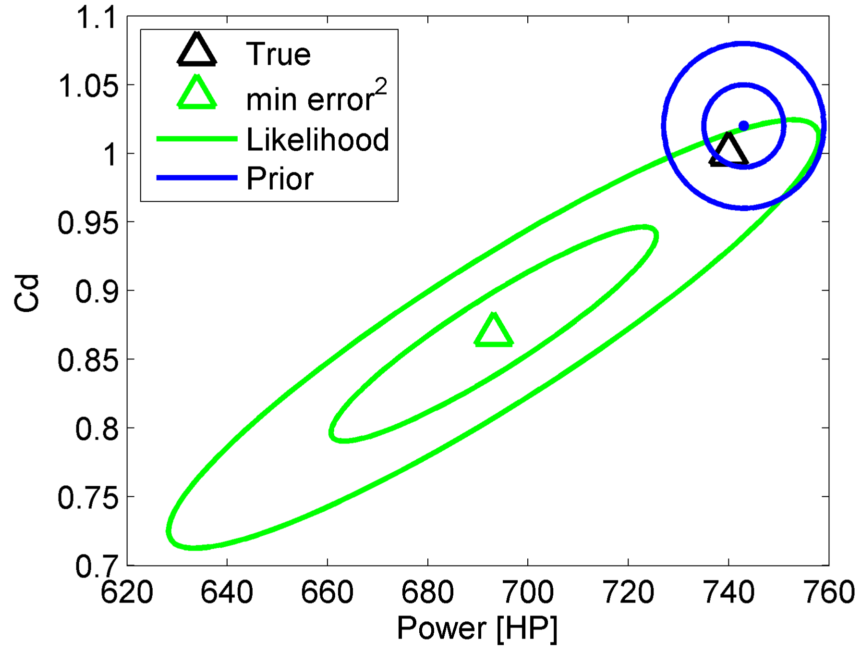
- Theoretical justification: if our model is *linear* and the noise in the measurements is *Gaussian*, minimising the sum of the squared errors is equivalent to finding the maximum likelihood estimator.

- Likelihood: how probable is the observed data for different settings of Power and Drag?



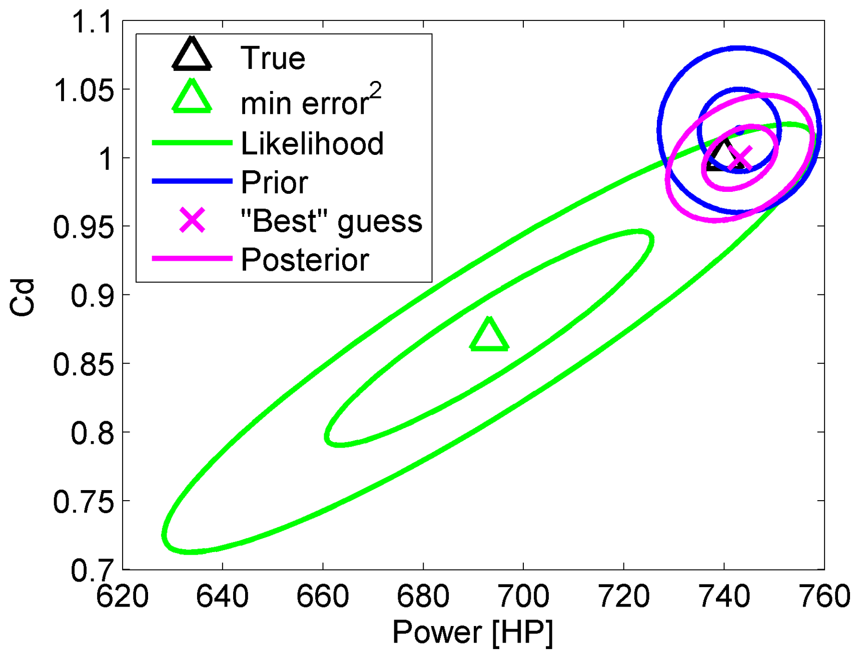
Beyond the Squared Errors

- Given what we already know about this car, we know that 693 HP and a Cd of 0.87 are both clearly too low.
- What levels of power and drag do we consider reasonable before seeing any data?
- In other words, what is our prior belief about the values of power and drag?



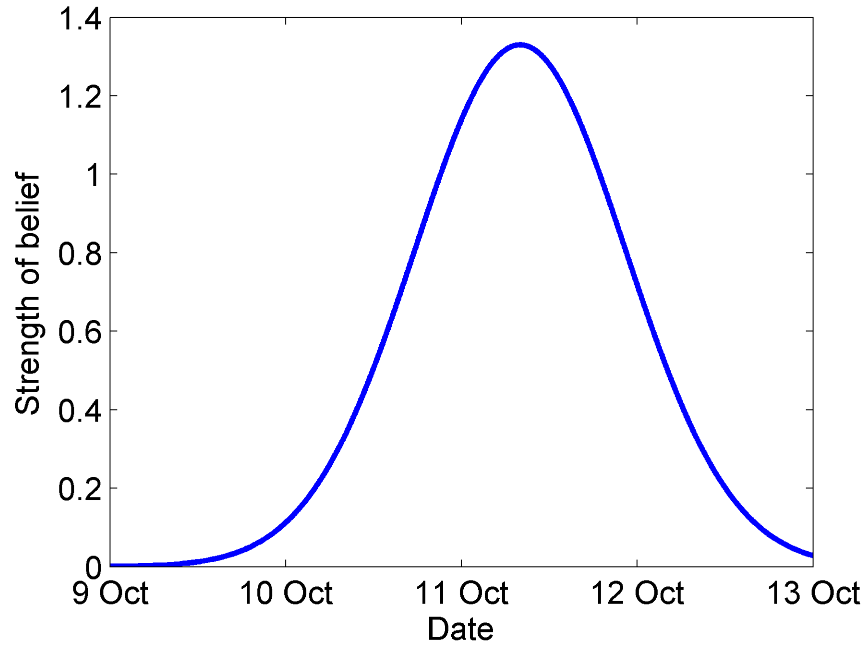
Beyond the Squared Errors

- We can combine our prior beliefs with the evidence provided by the observed data to compute the posterior: the uncertainty in (P,Cd) after we have observed the data.
- The posterior is not just a pair of numbers representing Power and Drag. It is a representation of our beliefs about the actual values of the Power and the Drag.



Modelling Uncertainty:
from Numbers to Probability Distributions

- We need a way to quantify degrees of belief: probability theory.
- We can consider it as an extension from logic (True,False) to a machinery that allows us to reason about uncertain statements that are between True and False.
- Example:

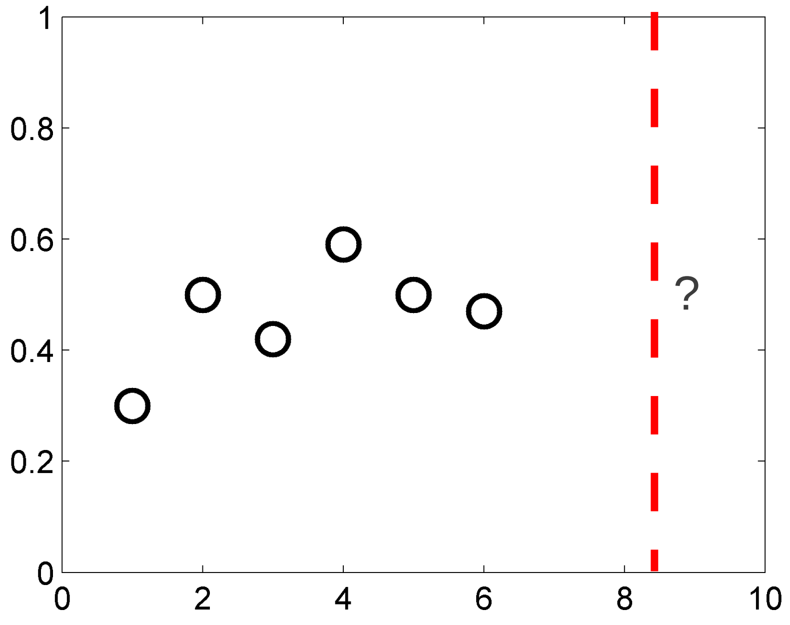


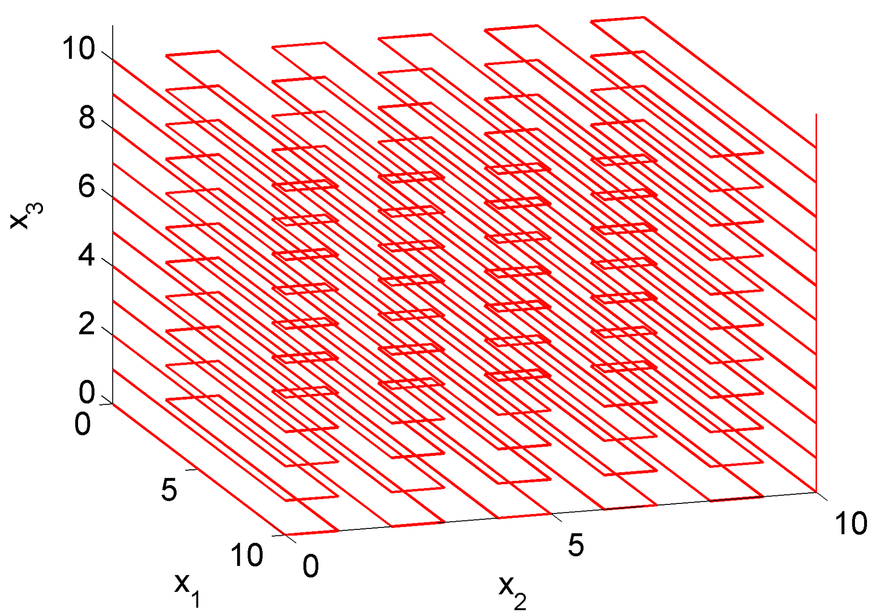
Standard Inference Problems

- Regression:
  - e.g. obtaining an AeroMap from a limited number of tunnel experiments.
- Classification:
  - e.g. finding which areas of an AeroMap are stalled
- Dimensionality Reduction:
  - e.g. finding a compact representation of the operating envelope of a car.
- Parameter Estimation:
  - e.g. estimating the value of tyre lateral stiffness.

Regression

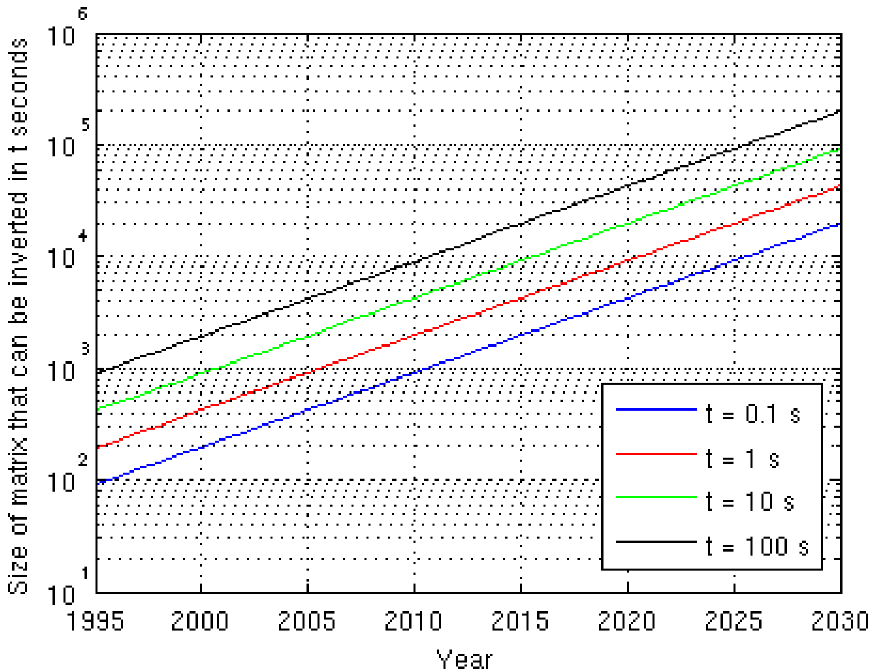
- We have measured lift and drag at a limited number of configurations:
  - What is the value of lift and drag in configurations we didn't test?
  - Given our measurements, what is the expected value of lap time with respect to the baseline?
- Regression and exploration are related.





So, should we model uncertainty everywhere?

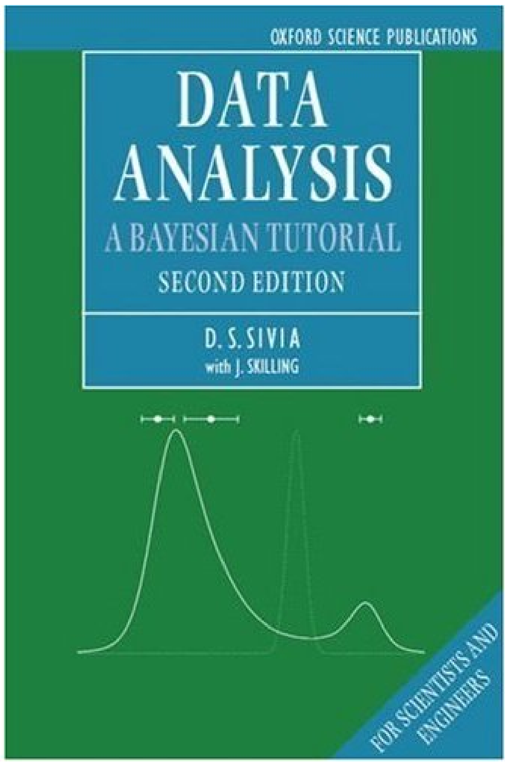
- Sometimes it is difficult!
- We may want to consider uncertainty just in some key areas.
- Some inference methods can be computationally demanding, but

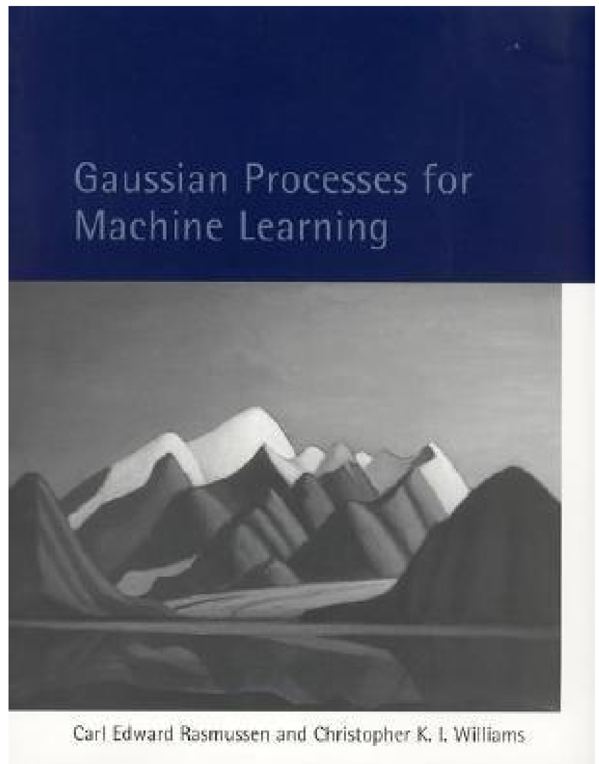


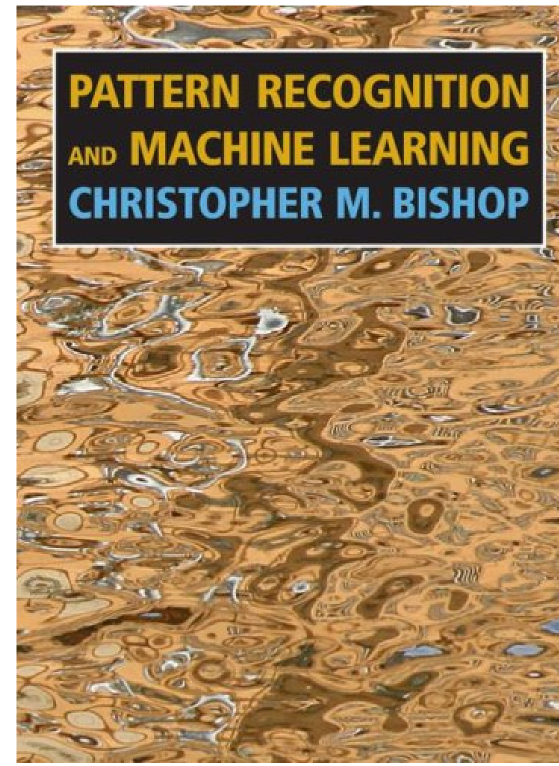
Canclusions

- “The applications are everywhere, because uncertainty is everywhere” Judea Pearl, Turing Award winner, 15 March 2012.
- Statistical inference is used extensively in areas where there is a lot of data and the need to make sense out of it: analysis of Internet data, genomics, speech recognition...
- If problems are formulated in a standard form we can then use many of the already developed methods.

Recommended Reading







Also, online Stanford courses: Machine Learning and Probabilistic Graphical Models