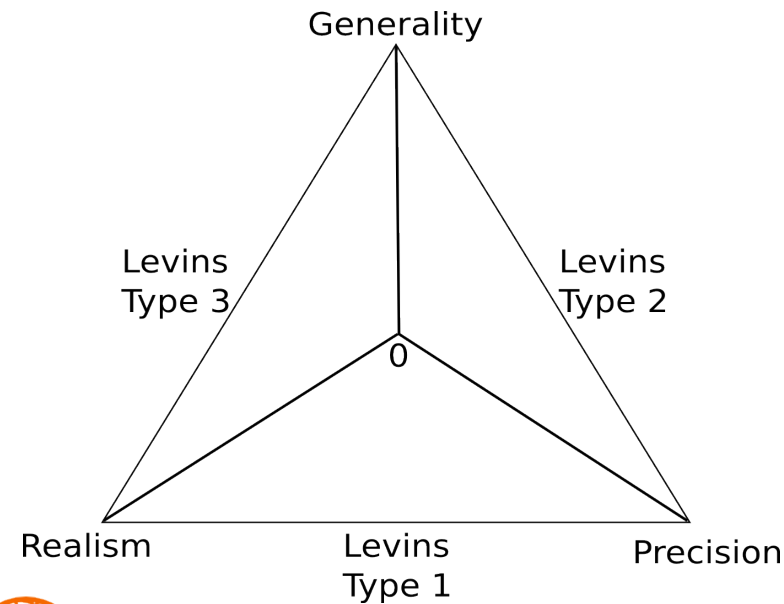
Realism in Statistical Analysis of Worst-Case Execution Times

David Griffin
Alan Burns

RTS York
THE UNIVERSITY of York

Levins and Model Building

- Levins (1966) proposed a system for categorising model building approaches



- Argues that no useful model can maximise the three desirable attributes: Generality, Realism and Precision
- Defined types of models which sacrifice one of these attributes

Levins applied to WCET

- Prediction, Estimation etc. all involve model building
- Existing techniques exhibit this tradeoff
  - e.g. Abstract interpretation is general and realistic, but not precise
  - Normally phrased as a tradeoff between one of these characteristics and tractability
    - Bullock and Silverman (2008) extended Levins argument to a fourfold tradeoff including tractability

Statistical Analysis

- Proposed by Edgar and Burns (2002)
- Uses Extreme Value Theory (EVT) Statistics to model execution times of a program
- Determines the probability with which a given deadline will be exceeded
  - When probability is low, other things break first...
- Refined by Hansen et al. (2009)
  - Usage closer to normal EVT usage
  - Produces failure rates

Statistical Analysis

- In terms of Levins, Statistical Analysis sacrifices realism
- When sacrificing realism, it's necessary to make sure that the model is realistic enough
- In the WCET problem, it's necessary to make sure that any sacrifice doesn't impact safety
- Edgar's experimental results had variable accuracy

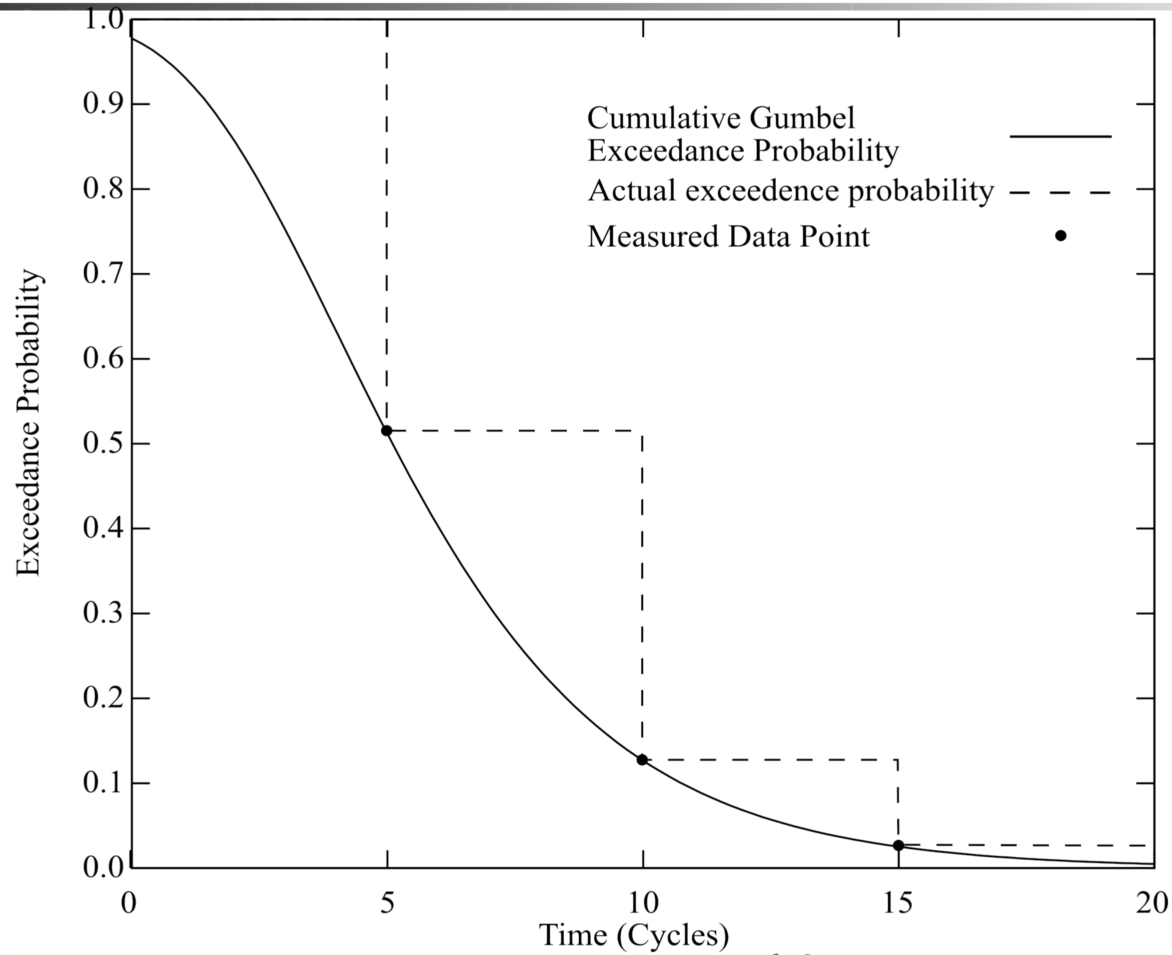
Decision Theory + Donald Rumsfeld

- There are...
  - Known Knowns: Things we know we know
  - Known Unknowns: Things we know we don't know
  - Unknown Unknowns: Things we don't know we don't know

Continuous vs Discrete Distributions

- The EVT Gumbel Distribution is Continuous
- Program runtimes are discrete
  - Processors use discrete time
  - Programs cannot terminate at any arbitary point
- Can unsafe errors be introduced by using EVT?

Continuous vs Discrete Distributions

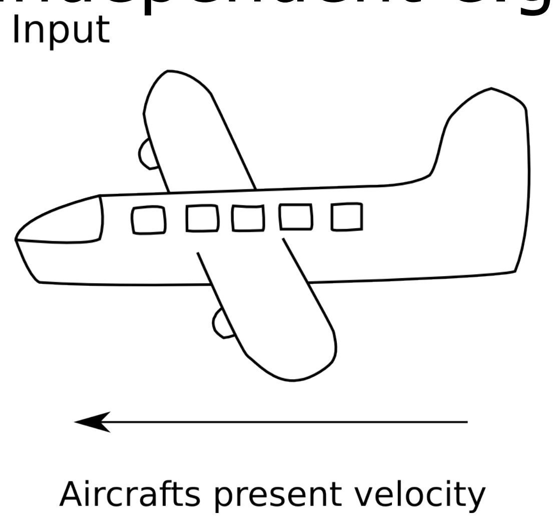


The I.I.D. Assumption

- EVT makes the i.i.d. Assumption
  - Independent: The probability of each outcome is not effected by outcomes which have already happened
  - Identically Distributed: The probability of each outcome is identical to the probability of the same outcome in another sample

The I.I.D. Assumption

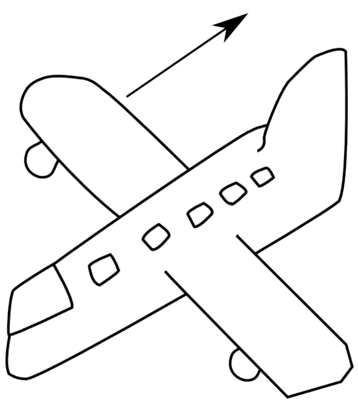
- Runtimes are not independent
- Processor caches in particular violate this
- Also some systems can never be independent e.g. Aircraft control



- Input: Current velocity of the aircraft
- Output: Modification to velocity of aircraft
- So input depends on previous output

The I.I.D. Assumption

- Runtimes are not independent
- Processor caches in particular violate this
- Also some systems can never be independent e.g. Aircraft control



- Input: Current velocity of the aircraft
- Output: Modification to velocity of aircraft
- So input depends on previous output

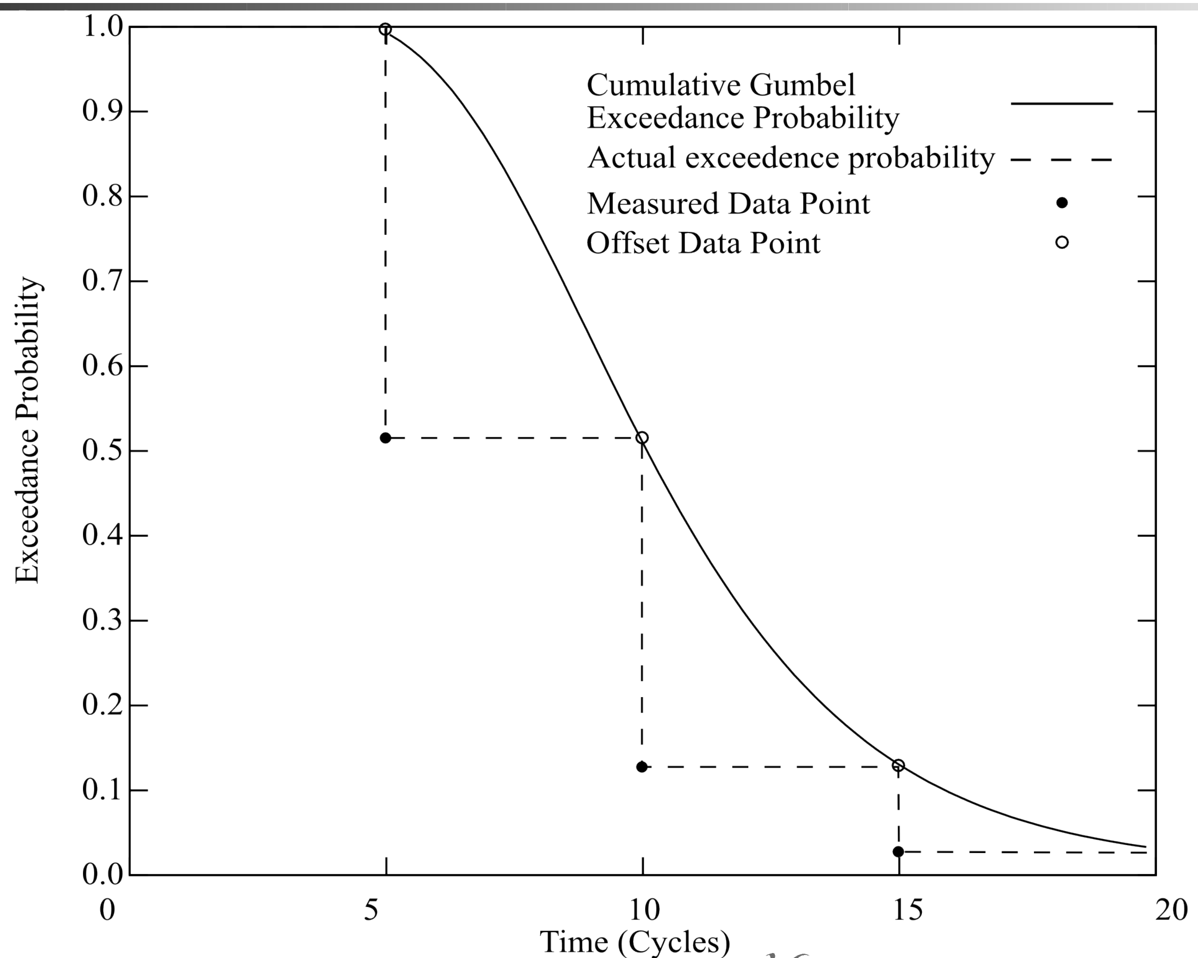
The I.I.D. Assumption

- Runtimes are not identically distributed
- On each path through the program, there are a number of hazards
- Separate paths through the program have different hazards
- So separate paths through the program have different distributions of runtimes
  - Whilst probability distributions can be joined, not all the distributions may be known

Compensating: Proof

- Argue that the problems don't apply or are bounded
- Not automatable, but some avenues to try
  - Independence: Statistical tests can give some confidence that dependence doesn't arise
  - Independence: Periodic resets to give a bound
  - Identically Distributed: Code coverage can give confidence that all distributions are found
  - Continuous approximation: Possible to modify the points being modelled to be safe

Compensating: Proof



Compensating: Adaption

- Change how Statistical Analysis is applied so it does not encounter problems
  - Identically Distributed: Use statistical analysis to explore one path through the program at a time
  - Independence: Perform resets / randomisation of shared state between tests
    - Not suitable for systems which must be dependent
  - Continuous approximation: Doesn't apply, as if exploring one path then large discrepencies cannot arise

Conclusions

- Statistical Analysis is potentially a very powerful tool
- But earlier work (Edgar and Burns (2002), Hansen et al. (2009)) does not guarantee that the results are safe
- For the results to be safe, either additional properties need to be proved or the method has to be applied in a more restricted form.