Institute of Flight System Dynamics

Technische Universität München

Predictive Flight Data Analysis

Lukas Höhndorf

Florian Holzapfel, Ludwig Drees, Javensius Sembiring, Chong Wang,
Phillip Koppitz, Stefan Schiele, Christopher Zaglauer

Institute of Flight System Dynamics
Technische Universität München
Garching, Germany

3rd Data Science in Aviation Workshop, Eurocontrol Headquarters
April 8th, 2015, Brussels, Belgium

Regulatory Framework

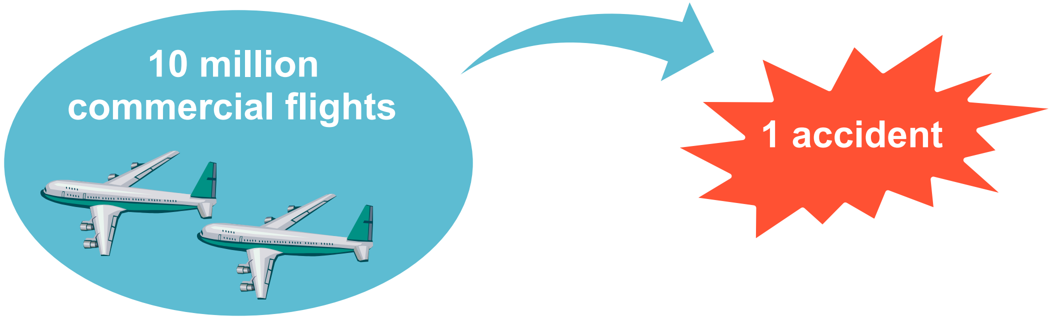
1

ICAO DOC 9859

- Airlines are required to implement a safety management system (SMS)
- SMS requires operators also to define their own **Acceptable Level of Safety (ALoS).**

"The minimum level of safety performance […] of a service provider, as defined in its safety management […] ."

- **Europe** aims at a **target accident rate** of less than one accident per ten million commercial flights (i.e. **accident probability of 10^{-7} per flight**).



2

Flightpath 2050

BUT: How to quantify the current level of safety?

Solution?

*ency of Incidents */ Number of Flights*

Classical statistical approach

?

$P_{Incident} = 0/400\,000 = 0$

Runway overrun example

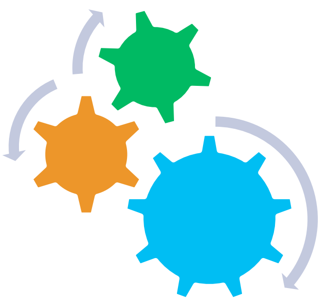
vs.

Classical statistical approach is inappropriate and **unsuitable for rare events**

* **Serious incidents** as defined in ICAO Annex 13

Mission Statement

- Predicting statistically valid accident probabilities for an individual airline based on available evidence from accident-free operation.
- Accounting for airline-specific factors such as operations, training, etc.



Predictive Analysis:
Making quantitative statements about the future state based on previous experience and knowledge.

BUT: **How to implement Predictive Analysis for practical application?**

Predictive Analysis:

Making quantitative statements about the future state based on:

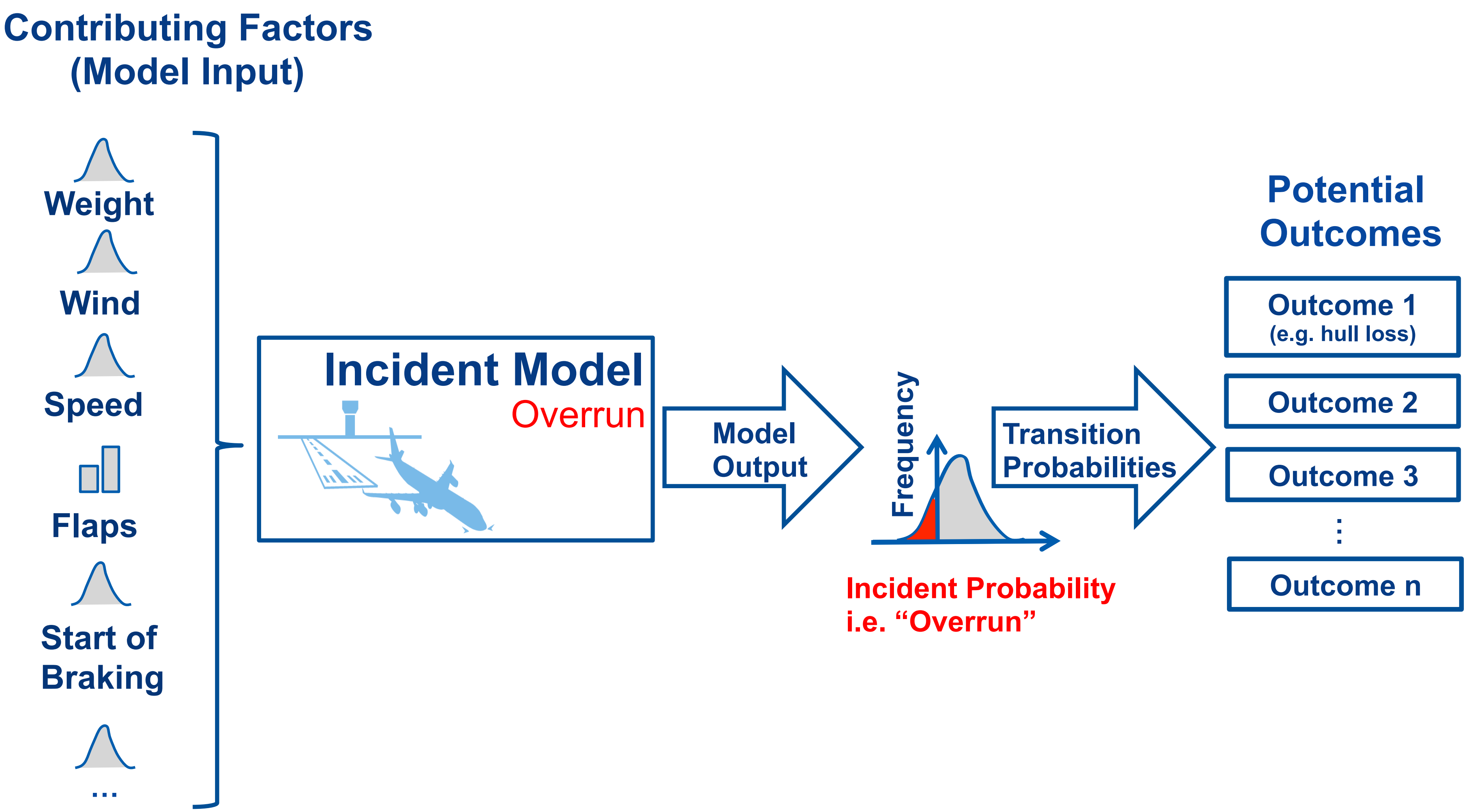
- previous experience
- knowledge

previous experience = data/evidence driven	• recorded data • known accident types and their causes
knowledge	• physical relation between contributing factors and accident • known cause-consequence-chains

Basic Hypothesis:

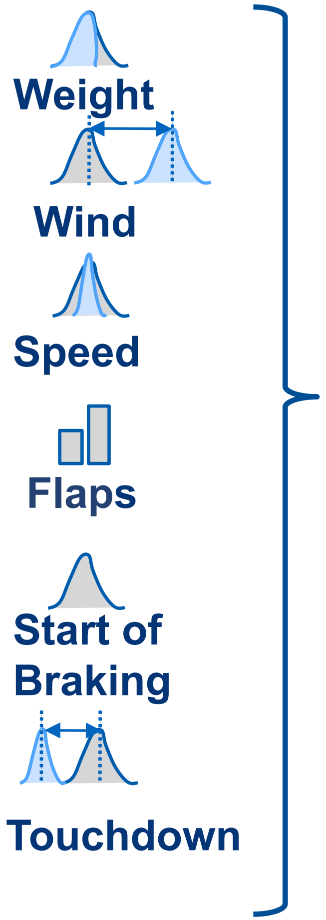
1. Accidents cannot be directly observed in daily operation, however, the contributing factors still occur at high frequency so they can be measured or observed with statistical significance.
2. The relation between the contributing factors and the accident can be described by the laws of physics and cause-consequence-chains based on operational and procedural knowledge.

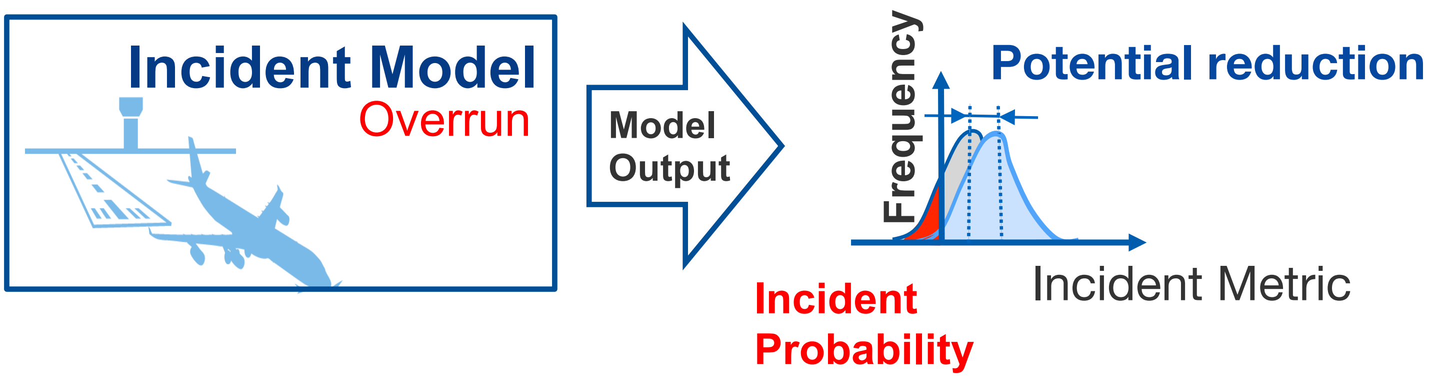
Predictive Analysis on Runway Overrun



Change Management

Contributing Factors (Model Input)





- Predictive analysis allows the assessment of the impact of mitigation actions **BEFORE** implementing them
- Impact of mitigation actions to **OTHER** incidents automatically considered (e.g. runway overrun vs. hard landing vs. tail strike)

British Airways BA038 Accident at London Heathrow

- Pilots were unable to increase speed during approach
- Boeing 777-236ER landed short of runway 27L
- Unknown dependencies between fuel flow and fuel temperature contributed to the accident [1]
- January 17^{th}, 2008

Source: http://news.bbc.co.uk

Source: http://www.thedigitalaviator.com

[1] AAIB Report on the accident to Boeing 777-236ER, G-YMMM, at London Heathrow Airport on 17 January 2008

Extract from the AAIB Report on the BA038 accident

"The investigation identified the following probable causal factors that led to the fuel flow restrictions:

1. Accreted ice from within the fuel system[1] released, causing a restriction to the engine fuel flow at the face of the FOHE, on both of the engines.

2. Ice had formed within the fuel system, from water that occurred naturally in the fuel, whilst the aircraft operated with low fuel flows over a long period and the localized fuel temperatures were in an area described as the "sticky range".

3. The **FOHE**, although **compliant with the applicable certification requirements**, was shown to be susceptible to restriction when presented with soft ice in a high concentration, with a fuel temperature that is below -10 °C and a fuel flow above flight idle.

4. Certification requirements, with which the aircraft and engine fuel systems had to comply, did not take account of this **phenomenon as the risk was unrecognized at that time**."

[1] For this report "fuel system" refers to the aircraft and engine fuel system upstream of the FOHE.

FOHE … Fuel Oil Heat Exchanger

Therefore our goal is to

Get a thorough description of dependencies between parameters relevant in terms of airlines safety management to discover HIDDEN influences!

Focus of attention and outlook

- Obtained information will be used for the predictive analysis in flight safety management
- Rare events and their dependencies
  Observe that the extreme and rare realizations contribute to an aircraft accident.

Correlation coefficient

Let X_1 and X_2 be two random variables with finite variances

$$corr(X_1, X_2) = Cov(X_1, X_2) / \sqrt{Var(X_1)} * \sqrt{Var(X_2)}$$

This is a measure of "LINEAR dependence" with range [-1,1], so this is ONE VALUE.

Some mathematical disadvantages

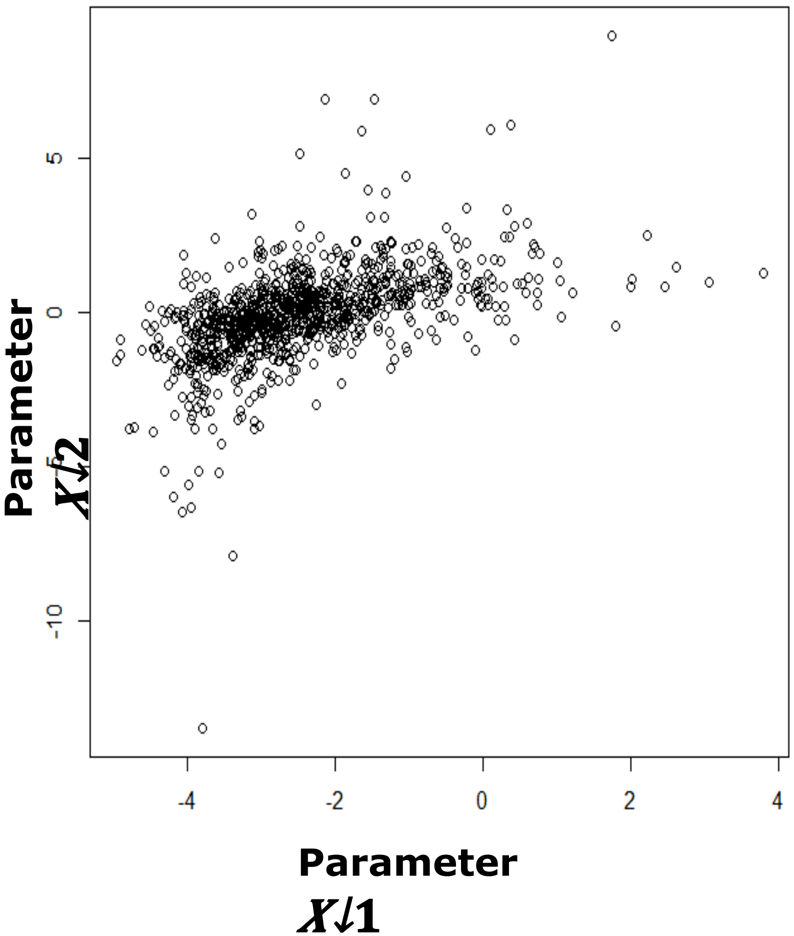
- Only defined for two random variables
- Higher dimensions cannot be represented simultaneously
- Non-linear dependencies are not captured properly

This is not a satisfying dependence measure for our application!

The concept of Copulas is more suitable.

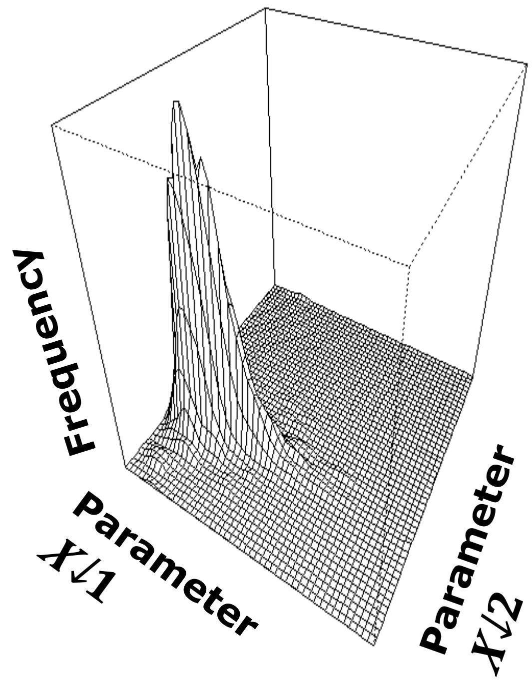
Dependence Modeling

- **Example Data:** Sample Size 1000



Obviously there is some kind of dependence between Values X and Values Y.

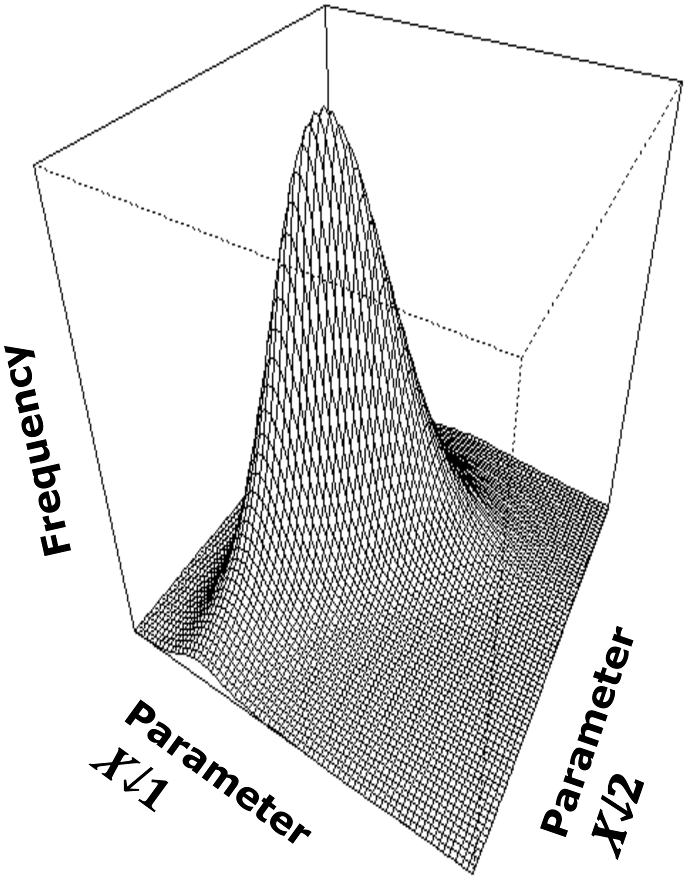
Laying a grid over the region and apply a Kernel Density estimation gives:



Investigation 1

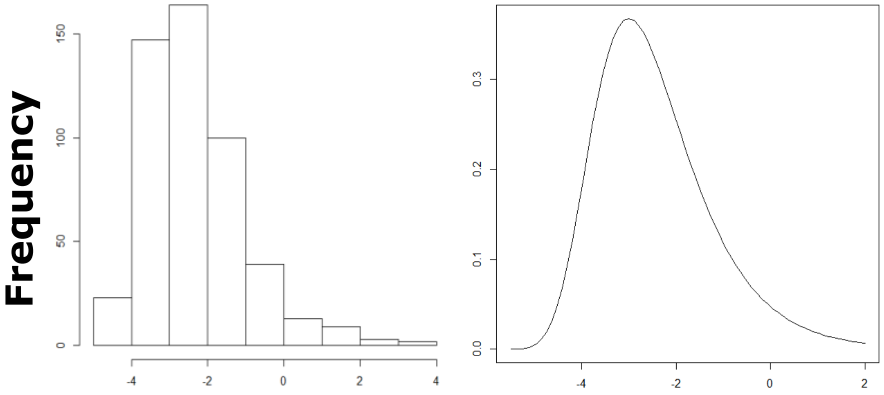
- Given the data we can estimate the "Joint Distribution".

- The estimation of the Joint Distribution for more than 3 Parameters is very difficult!

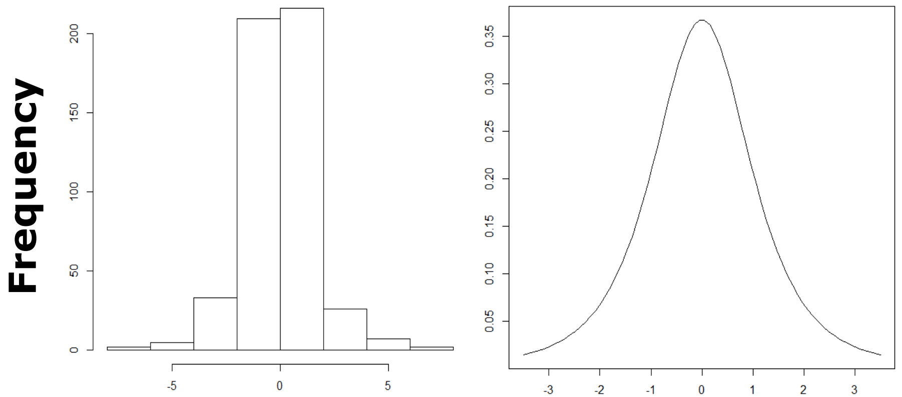


Investigation 2

- Alternatively we can concentrate on the distributions of the two values separately.
- The results are two “Marginal Distributions”.

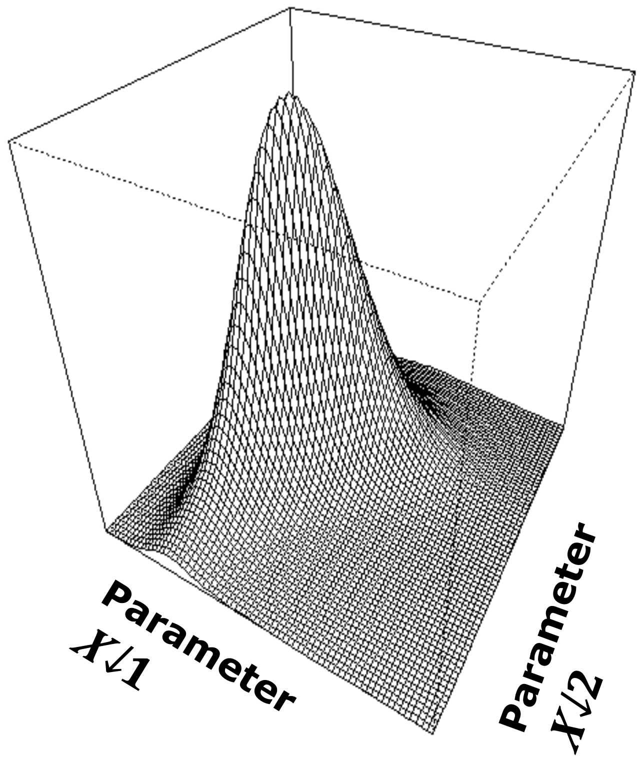


Parameter
X_1

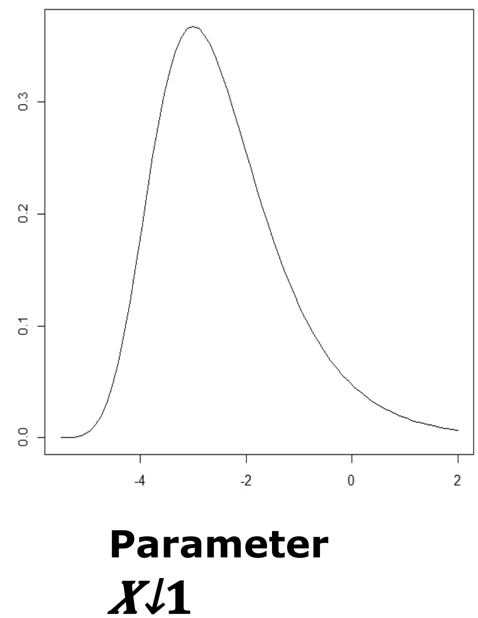


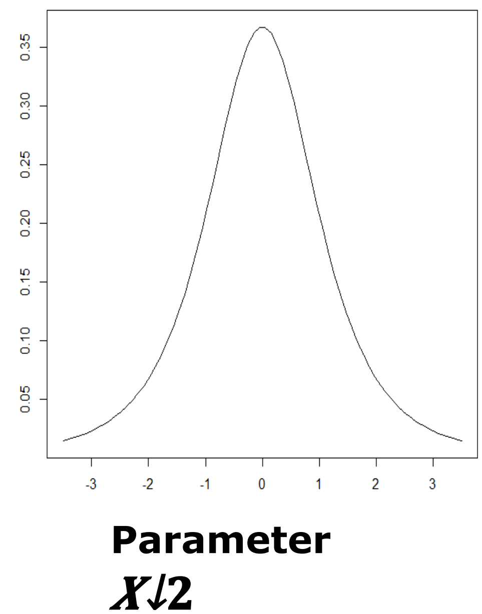
Parameter
X_2

Central Question: Which investigation gives more information?



Investigation 1 – Joint Distribution





Investigation 2 –
Two Marginal Distributions

- **Answer:** The Joint Distribution gives more information since the dependencies between the two parameters are included, but they are not represented within the two marginal distributions.

  (Consider: The marginal distribution can be calculated from the joint distribution by integration.)

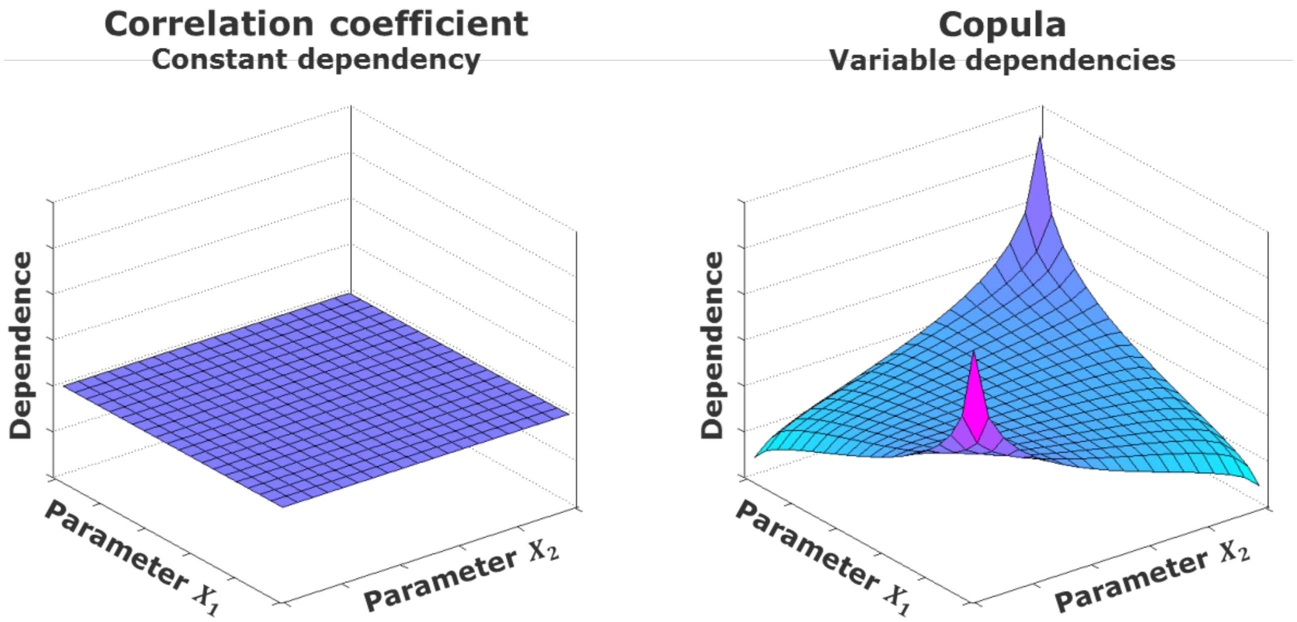
- But if we add a suitable Copula to the marginal distributions, the information is equal.

"d Marginals + 1 Copula = Joint Distribution in d dimensions"

Quantifying the dependence structure

- Simultaneous observation of several incidents is possible (e.g. Runway Overrun, Tailstrike and Hard Landing).

- The presented method might enable us to quantify unknown dependencies.



Tail Dependence Coefficients

Potential Hazard:

"Given that the average fuel temperature is small, what is the probability that the fuel flow is (too) small shortly ahead of landing?"

- For a bivariate distribution we define the (lower) tail dependence coefficient to evaluate the boundary behavior of dependence by setting

$$\lambda^{lower} := \lim_{t \to 0^+} P\left(X_2 \leq F_{X_2}^{-1}(t) \mid X_1 \leq F_{X_1}^{-1}(t)\right)$$

- In many practical cases this conditional probability might not be easy to calculate.

- With the Copula distribution function C we can calculate: $\lambda^{lower} = \lim_{t \to 0^+} C(t,t)/t$

SafeClouds – Big Data

Flight data

Obtained from the Quick Access Recorder

Weather data

Temporarily (for specific flight) and long-term (e.g. at the airport) investigation possible

ADS-B

To inspect for example TCAS related issues

Performance calculations

To compare estimated and real values.

Air traffic management information

The density of flights in a certain area can be considered

Terrain data

To evaluate recorded positions.

Technical data

Properties of technical devices (e.g. failure probabilities, cycles) or maintenance data

Data Mining

Professor

Florian Holzapfel (florian.holzapfel@tum.de)

Flight Safety Group

Ludwig Drees (ludwig.drees@tum.de)
Javensius Sembiring (javensius.sembiring@tum.de)
Lukas Höhndorf (lukas.hoehndorf@tum.de)
Chong Wang (chong.wang@tum.de)
Phillip Koppitz (phillip.koppitz@tum.de)
Stefan Schiele (stefan.schiele@tum.de)
Christopher Zaglauer (christopher.zaglauer@tum.de)

Thank you!