Analyzing uncertainty in civil engineering

Michael Oberguggenberger

Unit of Engineering Mathematics, Department of Engineering Science

University of Innsbruck, Austria

REC 2016, Ruhr University Bochum, June 15 – 17, 2016

June 15, 2016

Plan of Talk

- Discussion of uncertainty in civil engineering
- Models of uncertainty
- Semantics—interpretation
- Axiomatics—properties of aggregation
- Multivariate case
- Example: Winkler beam, hybrid models of uncertainty
- Sensitivity analysis
- Remarks

reality	$\longleftrightarrow$	**correspondence rules**	$\longleftrightarrow$	**model**
observations data		semantics interpretation		state variables parameters
$\uparrow$				$\downarrow$
action		$\longleftarrow$		computation

- *Definition and axiomatics:* How is uncertainty described and what are the combination rules?
- *Numerics:* How is uncertainty propagated through the computational model?
- *Semantics:* What is the meaning of the results – what do they say about our conception of reality?

Model uncertainties:

- choice of the structural model
- selection of state variables and parameters
- choice of limit state function
- transitional states during construction

Parameter uncertainties:

- random fluctuations
- lack of information
- random measurement errors
- systematic measurement errors
- fluctuations due to spatial inhomogeneity
- errors made by assigning parameter status to state variables
- parameters carry burden of model insufficiency

Range of available information:

- frequency distributions obtained from large samples
- values from small samples or single measurements
- interval bounds
- experts' point estimates
- educated guesses from experience

Failure probability

- R: all random variables describing the resistance of a structure
- S: all random variables describing the loads
- *Limit state function:* $g(R, S)$
- *Failure probability:* $p_f = P\big(g(R, S) < 0\big)$
- **Trouble** with p_f: Codes require $\mathbf{p_f = 10^{-6}}$

Imprecise Probability Models

Taking the stage in civil engineering in the 1990s:

Interval analysis, set-valued models, fuzzy sets, evidence theory, random sets, sets of probability measures, imprecise probability, lower and upper previsions, info-gap-analysis, etc.

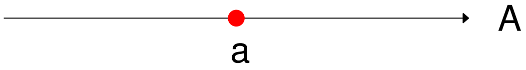
Understanding uncertainties:

- reflection about the choice of model and the failure mechanisms;
- assessing the variability of input and output variables and model parameters;
- sensitivity analysis;
- assessing the reliability of the structure;
- decision about acceptance or non-acceptance of the design.

> *Practical engineering = decision making with the help of scientific tools!*

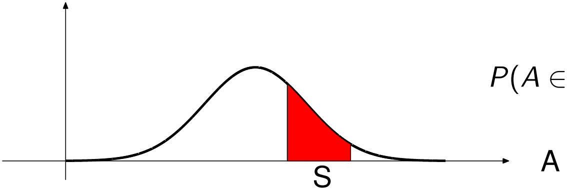
Models of Uncertainty—Definitions (1)

a. Deterministic description: The parameter A is described by a single value a.



b. Intervals: The uncertainty of the input A is described by an interval $[a_L, a_R]$.

c. Probability: The most informative, but also most stringent description of the uncertainty of a parameter A.



$$P(A \in S) = \int_S p_\lambda(a)\,\mathrm{d}a$$

The notation p_λ indicates that the probability distributions arise as members of a class of distributions, parametrized by parameters λ.

Example: $\mathcal{N}(\mu, \sigma^2)$ with $\lambda = (\mu, \sigma)$,

$$p_\lambda(a) = \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(a-\mu)^2}{2\sigma^2}}.$$

d. Sets of probability measures: Replace the single measure by a set of probability measures, a family

$$\mathcal{M} = \{p_\lambda : \lambda \in \Lambda\}.$$

A set of probability measures defines *lower and upper probabilities* according to the rules

$$\underline{P}(A \in S) = \inf\{P(A \in S) : P \in \mathcal{M}\},$$
$$\overline{P}(A \in S) = \sup\{P(A \in S) : P \in \mathcal{M}\}.$$

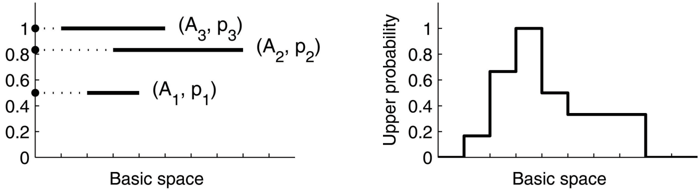
e. Finite random sets (Dempster-Shafer structures):

Focal elements: Subsets $A_i, i = 1, \ldots, n$ of a given set $\mathcal{A}$
Probability weights: $p_i = p(A_i)$, $\sum p_i = 1$.

Upper/lower probability (plausibility/belief) of an event S:

$$\overline{P}(S) = \sum_{A_i \cap S \neq \emptyset} p_i, \qquad \underline{P}(S) = \sum_{A_i \subset S} p_i$$

Contour function: $a \to \overline{P}(\{a\})$



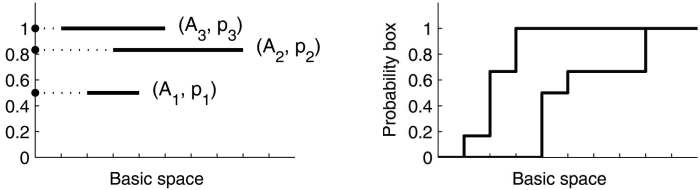
e. Finite random sets (Dempster-Shafer structures):

Focal elements: Subsets $A_i, i = 1, \ldots, n$ of a given set $\mathcal{A}$
Probability weights: $p_i = p(A_i)$, $\sum p_i = 1$.

Upper/lower probability (plausibility/belief) of an event S:

$$\overline{P}(S) = \sum_{A_i \cap S \neq \emptyset} p_i, \qquad \underline{P}(S) = \sum_{A_i \subset S} p_i$$

Probability box: $\overline{F}(b) = \overline{P}(-\infty, b]$, $\underline{F}(b) = \underline{P}(-\infty, b]$



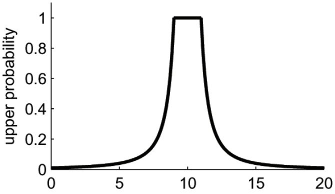
f. Tchebycheff random sets:

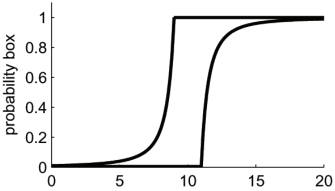
Nonparametric representation of random quantity with mean μ and variance σ^2, from

$$P\big(|a-\mu| > d(\alpha)\big) \le \alpha \quad \text{with} \quad d(\alpha) = \sigma/\sqrt{\alpha}$$

as uniformly distributed random set on $0 < \alpha \le 1$:

$$\alpha \to A(\alpha) = [\mu - d(\alpha), \mu + d(\alpha)]$$



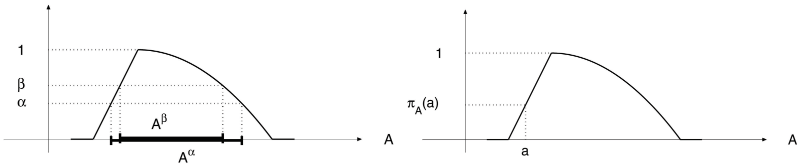


$\mu = 10$, $\sigma = 1$. Interpretation: piled confidence intervals.

g. Fuzzy sets: A fuzzy set A as a *family of parametrized intervals* $A^\beta \subset A^\alpha$, $0 \le \alpha \le \beta \le 1$ (left),

or as a *membership function*, assigning to each real number a a value $\pi_A(a) \in [0,1]$ (right).

Intervals correspond to α-*level sets* $A^\alpha = \{a : \pi_A(a) \ge \alpha\}$.



One can introduce a *possibility measure* on the real line, defining a degree of possibility for each subset by

$$\pi_A(S) = \sup\{\pi_A(a) : a \in S\}.$$

a. Probability: The most prevalent and important semantics in engineering practice are:

1. *Classical probability*, based on principles like the *principle of non-sufficient reason* would determine the probability of an event S as the fraction of favorable cases among the possible cases.
2. *Frequentist probability*, based on the idea of random occurrence of an event in a sequence of independent trials, would approximate the probability of an event S by its relative frequency.
3. *Subjective probability* is meant to be a measure of personal confidence. It can be assessed by introspection and/or elicitation through experts.

b. Sets of probability measures: Arising in the *frequentist setting* in the context of confidence and tolerance intervals (distribution parameters varying in intervals).

Robust statistics is based on distributions neighboring a given distribution.

In *Bayesian statistics*, families of prior distributions have been employed (robust Bayesian methods), as well as fuzzy prior distributions.

c. Random sets as sets of probability measures: Denote by $\mathcal{M}(A_i)$ the totality of all probability measures living on A_i.

The set of probability measures induced by the random set is

$$\mathcal{M} = \{P : P = \sum m(A_i)P_i, P_i \in \mathcal{M}(A_i)\}.$$

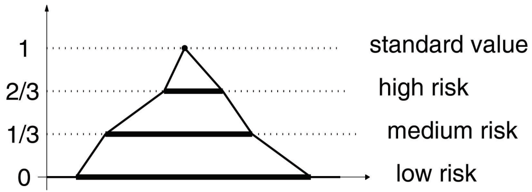
The respective notions of lower and upper probability coincide.

d. Fuzzy sets: The idea of *possibility* provides an interpretation for a fuzzy set and an operational method of constructing it.

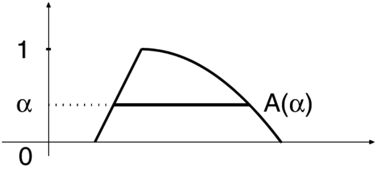
Introduce a scale $0 \le \alpha \le 1$ of risk levels specified verbally by the designing engineer.

Plot the ranges of fluctuation of the modeled parameter by intervals at their risk level.

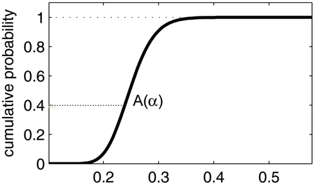
Join endpoints to obtain the membership function.



Every normalized fuzzy number can be seen as a random set; the sets $A(\alpha)$ are just the α-level sets.



Any random variable X can be reconstructed as a random set on $[0, 1]$ by putting $A(\alpha) = F_X^{-1}(\alpha)$.



The axioms fix the algebraic properties of the corresponding set functions.

Probability measures—additive set functions:

$$p(S \cup T) = p(S) + p(T) - p(S \cap T)\,.$$

Possibility measures—maxitive set functions:

$$\pi(S \cup T) = \max\{\pi(S), \pi(T)\}\,.$$

Plausibility measures (include probability and possibility):

$$\eta(S \cup T) \leq \eta(S) + \eta(T) - \eta(S \cap T)\,.$$

Monotone set functions (include all):

$$\mu(S \cup T) \geq \max\{\mu(S), \mu(T)\}\,.$$

Part of the Hierarchy

Focal elements singletons

Probability measures

$P(S \cap T) = P(S) + P(T) - P(S \cup T)$

Focal elements consonant

Possibility measures

$\pi(S \cup T) = \max\{\pi(S), \pi(S)\}$

Fuzzy sets

$\pi(a) = \pi(\{a\})$

$\pi(S) = \sup\{\pi(a): a \in S\}$

Focal elements arbitrary

Plausibility measures

$\eta(S \cap T) \leq \eta(S) + \eta(T) - \eta(S \cup T)$

Monotone set functions

$\mu(S \cup T) \geq \max\{\mu(S), \mu(S)\}$

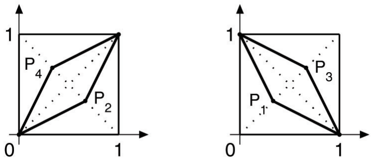
$A = (A_1, \ldots, A_n)$ ... in principle, as before.

But: Modeling mutual dependence, correlation, interaction.

Probability theory: Two random variables are *independent* when their joint distribution function is the product of the individual (marginal) distribution functions.

Interval analysis: Two parameters taking interval values are *non-interactive* when their joint range is the product interval (a rectangle), and *interactive* when their joint range is a proper subset of the product interval.

Fuzzy sets are non-interactive if their α-level sets are rectangles. Interactivity can be modeled by parametric restrictions on the α-level sets.



Positive/negative interactivity.

Random sets: The concept of independence splits up.

Random set independence: the joint focal sets are product intervals, and the joint weights are the products of the corresponding marginal weights.

Strong independence: The underlying set of joint probability measures consists of product measures only.

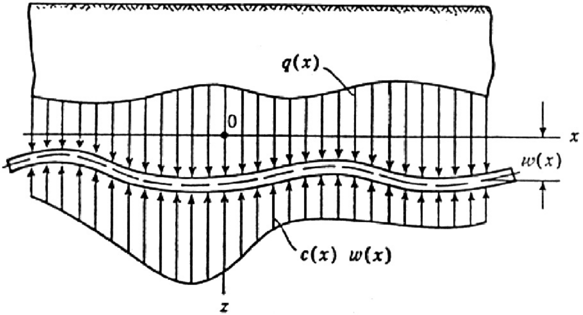
Fuzzy set independence: Product intervals from equal α-levels.

Modeling dependence by copulas: A copula $C(u, v)$ is a multivariate distribution function with uniform marginals.

Reconstructing a joint distribution $F_{XY}(x, y)$ of two random variables from the marginals:

$$F_{XY}(x, y) = C(F_X(x), F_Y(y)).$$

Application to random sets: Prescribing correlations on the basic probability weights by means of parametric copulas.



From: V. Bolotin, *Statistical Methods in Structural Mechanics*. San Francisco: Holden-Day 1969, Section 61.

Equation for the displacement u(x):

$$EI\, u^{IV}(x) + bc\, u(x) = q(x), \ -\infty < x < \infty\,,$$

EI . . . flexural rigidity of beam
b . . . effective width
c . . . bearing coefficient of foundation
$q(x)$ . . . loading

Standardized equation:

$$u^{IV}(x) + 4k^4 u(x) = p(x), \; -\infty < x < \infty$$

with $bc/EI = 4k^4$, $p(x) = q(x)/EI$.

Unique deterministic solution:

$$u(x) = \frac{1}{8k^3} \int_{-\infty}^{\infty} e^{-k|x-y|} \big(\sin k|x-y| + \cos k|x-y| \big)\, p(y)\, dy$$

This is the I-O-map used in Monte Carlo simulations.

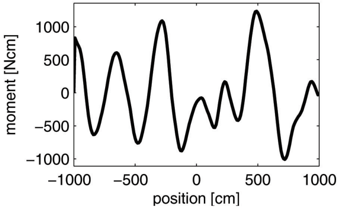
Deterministic reference case:

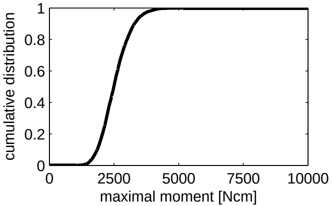
$q(x) \equiv q = 10$ [N/cm], $\mathrm{EI} = 10^9$ [Ncm^3],
$b = 6$ [cm], $c = 6.7$ [N/cm^3], whence $bc \approx 40$ [N/cm^2].

Assumption: $bc = 40$ [N/cm^2] deterministic; load as a random field $q(x, \alpha)$ with mean $\mu_q = 10$, field variance $\sigma_q^2 = 4$, correlation length $\ell = 1000$ [cm]. Autocorrelation function:

$$\mathrm{corr}(\rho) = \exp\big(-|\rho|/\ell\big).$$

Computation of maximal bending moment by Monte Carlo simulation, $N = 5000$ trajectories, $M_{\max} = \max(EIu''(x))$:





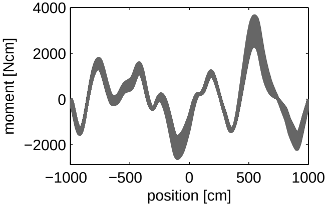
Bending moment (trajectory), maximal bending moment (simulated cumulative distribution).

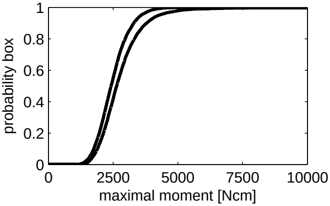
Load as a random field $q(x, \alpha)$ as above, coefficient bc as a Gaussian random variable with mean $\mu_{bc} = 40$ and interval-valued standard deviation $[\underline{\sigma}_{bc}, \overline{\sigma}_{bc}] = [4, 12]$.

(This results from a spread of $\pm 50\%$ around the nominal value $\sigma_{bc} = 8$ deriving from a coefficient of variation of 20%.)

Solution – the set-valued stochastic process

$$U(x, \alpha) = \{u(x, \alpha, \sigma_{bc}) : \sigma_{bc} \in [4, 12]\}.$$





Bending moment (interval trajectory), maximal bending moment (p-box).

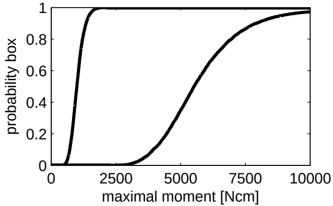
Load $q(x, \alpha)$ as set-valued random field, coefficient bc as imprecise Gaussian random variable:

Fixed:

$$\mu_{bc} = 40, \quad \mu_q = 10$$

Intervals:

$$[\underline{\sigma}_{bc}, \overline{\sigma}_{bc}] = [4, 12], \quad [\underline{\sigma}_q, \overline{\sigma}_q] = [1, 3], \quad [\underline{\ell}, \overline{\ell}] = [500, 1500]$$



P-box for maximal bending moment with with imprecise bc and q.

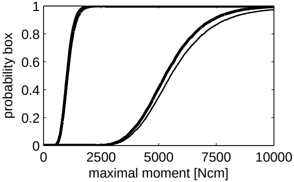
- Determining the degree of influence of individual input variables on the output
- Sampling based methods
- Successive pinching of variables and observing change of variability of output
- Random set approach using p-box requires no additional computational cost
- Quantification by information theoretic concepts, e.g. the Hartley-like measure

$$\mathrm{HL} = \int_0^1 \log_2 \left(1 + \underline{F}^{-1}(\alpha) - \overline{F}^{-1}(\alpha)\right) \mathrm{d}\alpha$$

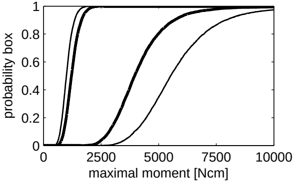
applied to the p-box.

This follows a suggestion of George Klir.

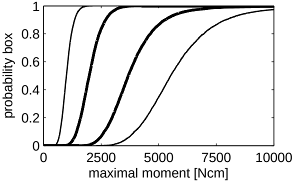
σ_{bc} pinched, σ_q and ℓ intervals



ℓ pinched, σ_{bc} and σ_q intervals



σ_q pinched, σ_{bc} and ℓ intervals



Choice of model, depending on available information:

Single variable

available information	model of uncertainty
upper/lower bounds	interval
value of mean and variance	Tchebycheff random set
type, value of mean and variance	probability distribution
type, bounds for mean and variance	imprecise probability distribution

Field variable

available information	model of uncertainty
type, value of mean, variance, correlation	random field
type, bounds for mean, variance, correlation	set-valued random field

- **Random sets as unifying framework**
- **Probability and intervals on various levels**
- **Numerics: Monte Carlo simulation, reweighting, surrogate models**
- **Alternative: sets of probability measures; gives tighter bounds**