Workshop on Risk Acceptance and Risk Communication
Stanford University, March 26-27, 2007

Aleatory or epistemic? Does it matter?

Armen Der Kiureghian
University of California, Berkeley

Ove Ditlevsen
Technical University of Denmark

An old story, retold …

Definitions:

Aleatory: From Latin *alea* = rolling of dice; uncertainty that arises from intrinsic randomness of a phenomenon.

Epistemic: From Greek $\varepsilon\pi\iota\sigma\tau\eta\mu\eta$ (*episteme*) = knowledge; uncertainty that arises from lack of knowledge (or data).

Generally accepted criterion:

- If the uncertainty can be reduced by increasing knowledge or gathering of more data $\Rightarrow$ epistemic
- If the uncertainty cannot be reduced by increasing knowledge or gathering of more data $\Rightarrow$ aleatory

Element of risk analysis

$\mathbf{X} = (X_1, \ldots, X_n)$ vector of basic random variables

$f_{\mathbf{X}}(\mathbf{x}, \mathbf{\Theta}_f)$ distribution of $\mathbf{X}$

$y_i = g_i(\mathbf{x}, \mathbf{\Theta}_g), \quad i = 1, \ldots, m$ predictive physical models

$\mathbf{\Theta}_f$ parameters in probabilistic model

$\mathbf{\Theta}_g$ parameters in physical models

Sources of uncertainty

- Uncertainty inherent in the basic random variables $\mathbf{X}$
- Uncertain error in the form of the probabilistic model $f_{\mathbf{X}}(\mathbf{x},\mathbf{\Theta}_f)$
- Uncertain errors in the physical models $y_i = g_i(\mathbf{x},\mathbf{\Theta}_g)$, $i = 1, \ldots, m$
- Statistical uncertainty in estimation of distribution parameters $\mathbf{\Theta}_f$
- Statistical uncertainty in estimation of physical model parameters $\mathbf{\Theta}_g$
- Uncertain measurement errors
- Uncertainty inherent in derived variables $\mathbf{Y}$
- Human error

Nature of uncertainties:
Uncertainty in basic random variables $\mathbf{X}$

Can be epistemic or aleatory:

e.g., uncertainty inherent in the strength of concrete

existing building $\Rightarrow$ epistemic

future building $\Rightarrow$ aleatory

There is a fundamental difference in reliability analysis of existing and future structures.

Nature of uncertainties:
Physical model uncertainty

$y = g(\mathbf{x}, \mathbf{z})$ exact but unknown relation

$y = \hat{g}(\mathbf{x}, \mathbf{\Theta}_g) + \varepsilon$ predictive model, $\varepsilon = N(0, \sigma_\varepsilon)$

Two components:

- Error due to missing variables $\mathbf{z}$ (can be aleatory or epistemic)
- Error in model form $\hat{g}$ (mostly epistemic)

Nature of uncertainties:
Probabilistic model uncertainty

$f_{\mathbf{X}}(\mathbf{x}, \mathbf{\Theta}_f)$ fitted to observed data or based on a priori assumptions

- Small probabilities are sensitive to tails of distributions
  $\Rightarrow$ uncertain error of epistemic type, difficult to assess
- Standard goodness-of-fit tests do not guarantee fit in the tail (Ditlevsen 1994).
- Need for standardization of distribution models in codes
- Caution must be exercised in performance-based engineering applications, which rely on absolute probability estimates

Nature of uncertainties:
Parameter uncertainty

$\mathbf{\Theta}_f, \mathbf{\Theta}_g$ estimated based on statistical analysis of observed data $\Rightarrow$ statistical uncertainty

$f_{\mathbf{\Theta}}(\boldsymbol{\theta})$ posterior distribution

- All statistical uncertainty is epistemic in nature.

Nature of uncertainties:
Remarks

- The choice between basic and derived variables depends on the available data and models.
- As scientific knowledge advances, more uncertainties that appeared to be aleatory become epistemic in nature.
- Why then differentiate? Differentiating aleatory and epistemic uncertainties helps us identify areas, where uncertainty can be reduced and models can be improved in near term. It also helps us to more accurately formulate risk and reliability problems.

Influence of uncertainties

Proper understanding of the nature of uncertainties is essential for formulation of risk and reliability problems.

Two demonstrative examples:

- Statistical dependence among system components due to statistical uncertainty.
- Statistical dependence among successive events in time due to non-ergodic uncertainties.

Example 1: System reliability

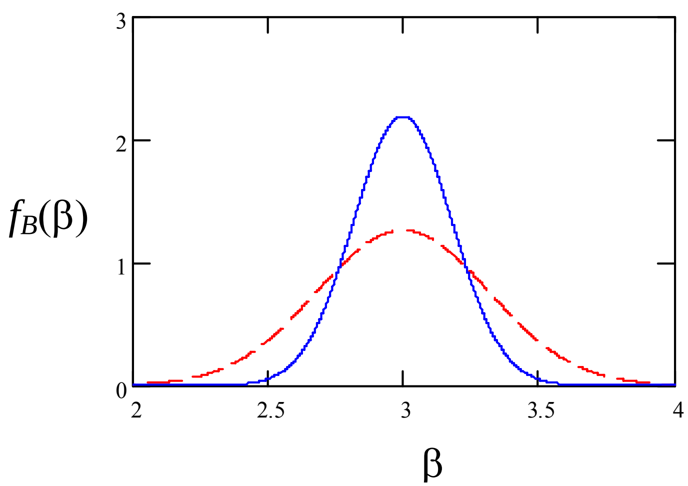
k-out-of-*N* system with statistically independent and identically distributed components:

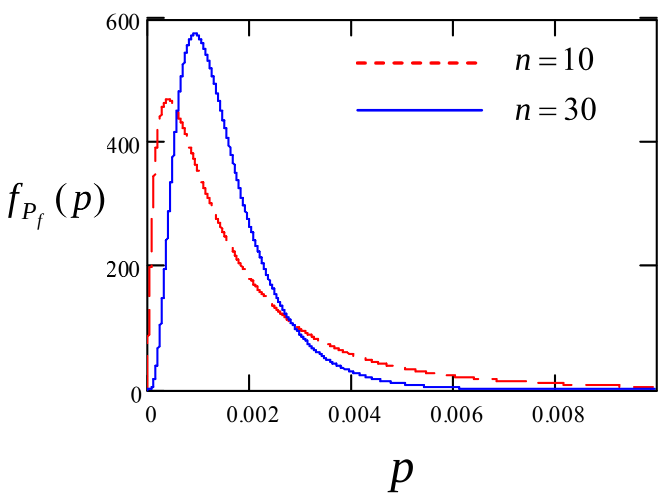
$g(\mathbf{x}) = x_1 - x_2$ limit-state for typical component

$X_i = \mathrm{N}(M_i,\sigma_i)$, σ_i known, $M_i = \mathrm{N}(\bar{x}_i,\sigma_i/\sqrt{n})$, $i = 1,2$

$B = \beta(M_1,M_2)$ Bayesian uncertain component reliability index

$P_f = \Phi^{-1}[\beta(M_1,M_2)]$ Bayesian uncertain component failure probability





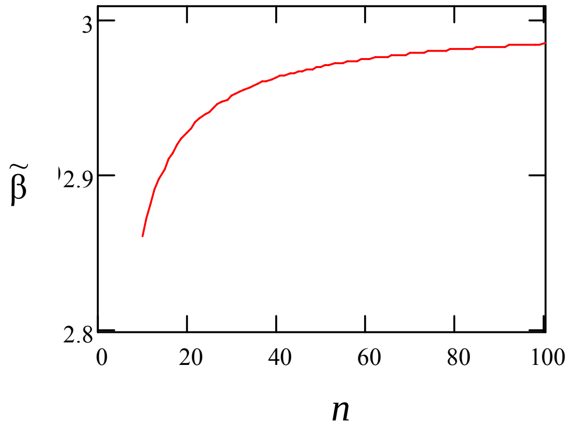
Example 1: System reliability

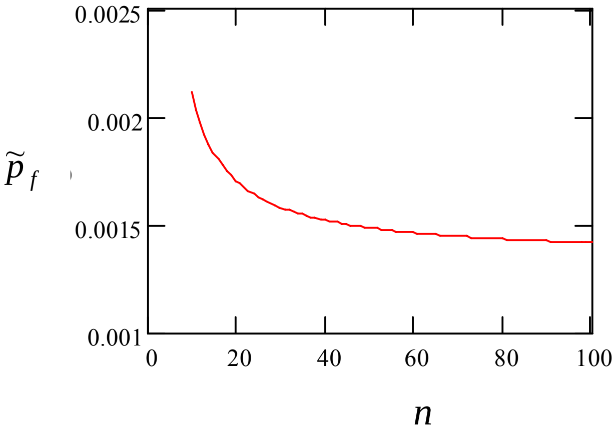
k-out-of-*N* system with statistically independent and identically distributed components:

Predictive component failure probability:

$$\tilde{p}_f = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p_f(\mu_1,\mu_2)\, f_{M_1}(\mu_1)\, f_{M_2}(\mu_2)\, \mathrm{d}\mu_1\, \mathrm{d}\mu_2$$

$$= \int_0^1 p f_{p_f}(p)\, \mathrm{d}p$$



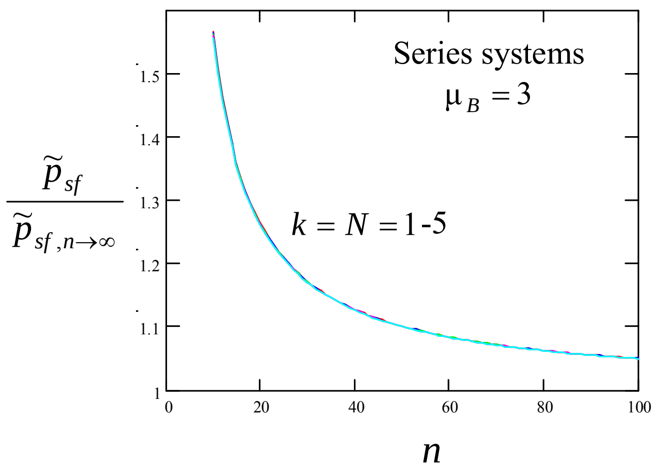


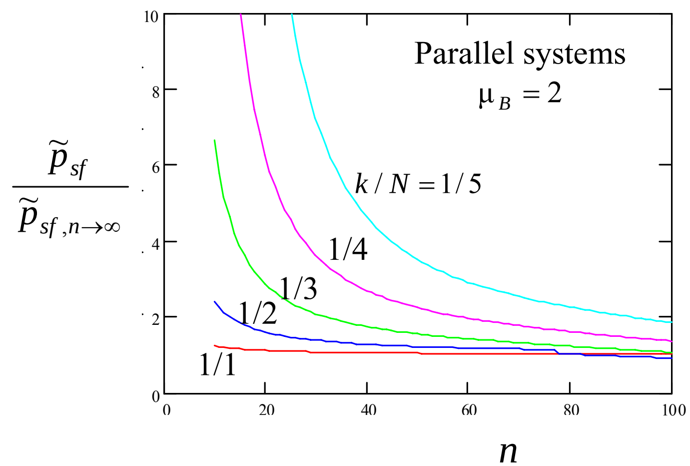
Example 1: System reliability

k-out-of-N system with statistically independent and identically distributed components:

Predictive system failure probability:

$$\tilde{p}_{sf} = \int_0^1 \sum_{j=N-k+1}^{N} \frac{N!}{j!(N-j)!} p^{\,j} (1-p)^{N-j} f_{P_f}(p)\,\mathrm{d}p$$



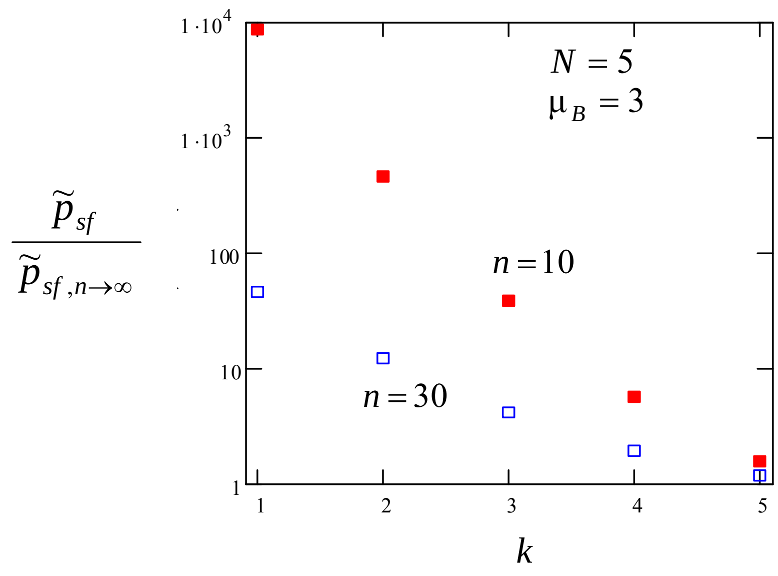


Example 1: System reliability

k-out-of-*N* system with statistically independent and identically distributed components:

Predictive system failure probability:

$$\tilde{p}_{sf} = \int_0^1 \sum_{j=N-k+1}^{N} \frac{N!}{j!(N-j)!} p^{j} (1-p)^{N-j} f_{P_f}(p)\,\mathrm{d}p$$



Example 2: Time-variant reliability

Structure subject to repeated earthquake loads modeled as Poisson events with mean rate ν per year:

Predictive system failure probability:

$R = LN(\lambda_R, \zeta_R)$ structure capacity

$S = LN(\lambda_S, \zeta_S)$ earthquake load at each occurrence

$g(r, s, \varepsilon) = \ln r + \varepsilon_1 - \ln s + \varepsilon_2$ limit-state model

$\varepsilon_1 = N(0, \sigma_1)$ capacity model error

$\varepsilon_2 = N(0, \sigma_2)$ load model error

$\Theta = (\nu, \lambda_R, \zeta_R, \lambda_S, \zeta_S, \sigma_1, \sigma_2)$ model parameters

Example 1: Time-variant reliability

Structure subject to repeated earthquake loads modeled as Poisson events with mean rate ν per year:

$$\tilde{p}_f = \int_{\boldsymbol{\theta}} \Phi\left[-\frac{\lambda_R - \lambda_S}{\sqrt{\zeta_R^2 + \sigma_1^2 + \zeta_S^2 + \sigma_2^2}} \right] f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}$$

predictive failure probability at each occurrence

$$\tilde{P}_{f,Psn} = 1 - \exp\left(-\mu_\nu\, \tilde{p}_f\, t\right)$$

predictive failure probability if events are assumed to be Poisson (the conventional approach)

$$\tilde{P}_f = 1 - \int_{r,e,\boldsymbol{\theta}} \exp\left[-\nu \Phi\left(-\frac{\ln r + e - \lambda_S}{\sqrt{\zeta_S^2 + \sigma_2^2}} \right) t \right] f_R(r)\, f_{\varepsilon_1}(e)\, f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})\,\mathrm{d}r\,\mathrm{d}e\,\mathrm{d}\boldsymbol{\theta}$$

predictive failure probability, correct solution

Example 1: Time-variant reliability

Structure subject to repeated earthquake loads modeled as Poisson events with mean rate ν per year:

Assumed parameter values:

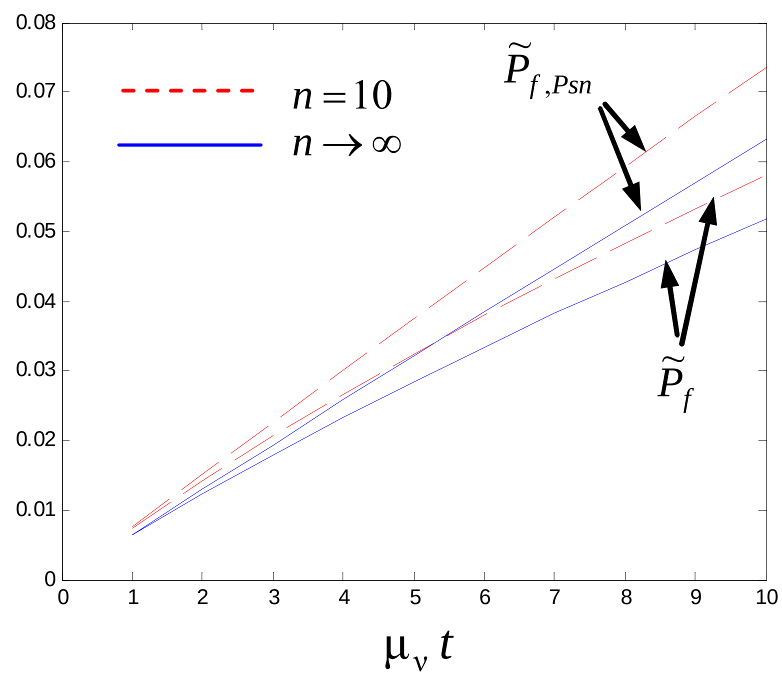
$$\nu = LN(\text{mean} = \mu_\nu, \text{c.o.v.} = 0.5)$$

$$\zeta_R = 0.294, \quad \lambda_R = N(0, \zeta_R / \sqrt{n}),$$

$$\zeta_S = 0.472, \quad \lambda_S = N(-1, \zeta_S / \sqrt{n}),$$

$$\sigma_1 = 0.3$$

$$\sigma_2 = 0.5$$



Summary and conclusions

- The nature of uncertainties (aleatoric or epistemic) depends on circumstances and modeling assumptions.
- Perhaps in the final analysis, all uncertainties are epistemic. However, characterization as aleatory or epistemic helps us in identifying uncertainties that can be reduced in short term by improving our models and by collecting data.
- Characterization of uncertainties is also important for proper formulation of risk and reliability problems. In particular, epistemic uncertainties can introduce correlation between the components of a system, and both types of uncertainties can introduce dependence among successive occurrences of events in time.