Baraldi Piero

Zio Enrico

Compare Michele
Energy Department, Polytecnic of Milan, Milan, Italy

Importance measures in presence of uncertainties

Keywords

importance measures, birnbaum importance, uncertainty, sorting algorithm, quicksort

Abstract

This paper presents a work on the study of importance measures in presence of uncertainties originating from the lack of knowledge and information on the system (epistemic uncertainties). A criterion is proposed for ranking the risk contributors in presence of uncertainties described by probability density functions.

1. Introduction

Importance Measures (IMs) are used to rank the contributions of components or basic events to the system performance, which can be, for example, the system reliability or risk. IMs were initially introduced by Birnbaum [1] in order to assess the contribution of the components to the overall system reliability; later different IMs have been introduced to address various aspects of reliability, safety and risk significance (Fussel-Vesely, Criticality, Risk Achievement Worth and Risk Reduction Worth) [2].

On the other hand, uncertainties of two types affect the behavior of a system [3]: aleatory and epistemic; the former type (also referred to as irreducible or stochastic or random uncertainty) describes the inherent variation associated with the physical system or the environment (e.g. variation in atmospheric conditions, in fatigue life of compressor and turbine blades); the latter (also referred to as subjective or reducible uncertainty) is, instead, due to a lack of knowledge of quantities or processes of the system or the environment (e.g. lack of experimental data to characterize new materials and processes, poor understanding of coupled physics phenomena, poor understanding of accident initiating events).

In practice IMs are calculated without due account of uncertainties. The objective of this work is then to investigate how uncertainties can influence IMs and how they can be accounted for. The uncertainties considered are of epistemic type and represented by probability density functions. A method is proposed for ranking the contributors to the system performance measure.

The paper is organized as follows. In Section 2, the problem of comparing the importance measures of two components whose reliabilities are uncertain is presented to explain the idea beyond the ranking method. In Section 3, two case studies are described: the first applies the comparison method on the components of a simple system, when uncertainties affect their reliabilities; the second validates the method on a large system for which a procedure is introduced for efficiently performing the ranking. Some conclusions are provided with regards to the comparison between the proposed procedure and a method previously presented in the literature [4].

2. Comparing the importance of two components in presence of uncertainties.

The aim of this Section is to present a procedure for comparing the importance of two components A and B of an hypothetical system in presence of uncertainty. In general, with respect to the uncertainty representation, when sufficiently informative data are available probabilistic distributions are righteously used. The uncertainties associated with the components performance characteristics should then be propagated through the system model, leading to uncertainties in the system performance. Hence, when performing importance measures calculations in presence of uncertainties affecting the components performances, the results should properly account for such uncertainties and so should the ranking.

For simplicity, uniformly distributed uncertainty is assumed to be affecting directly the IMs of components A and B. *Table 1* reports the range of the IMs distributions while *Figures 1* a and b show the corresponding distributions. The IM of component B is significantly more uncertain than that of component A.

Table 1. Uniform distributions parameters.

	Uniform distribution	
	Lower limit a	upper limit b
A	0.0141	0.0155
B	0.002	0.0178

Looking at the distributions of A and B IMs ($\mathbf{I_A}$ and $\mathbf{I_B}$) one may observe that E[$\mathbf{I_A}$] is greater than E[$\mathbf{I_B}$]; on the other hand there is a range in which the $\mathbf{I_B}$ quantiles are larger than the $\mathbf{I_A}$ ones. For example, if one were to perform the ranking based on the IMs 95th quantile values, the conclusion would be that component B is more important than A.

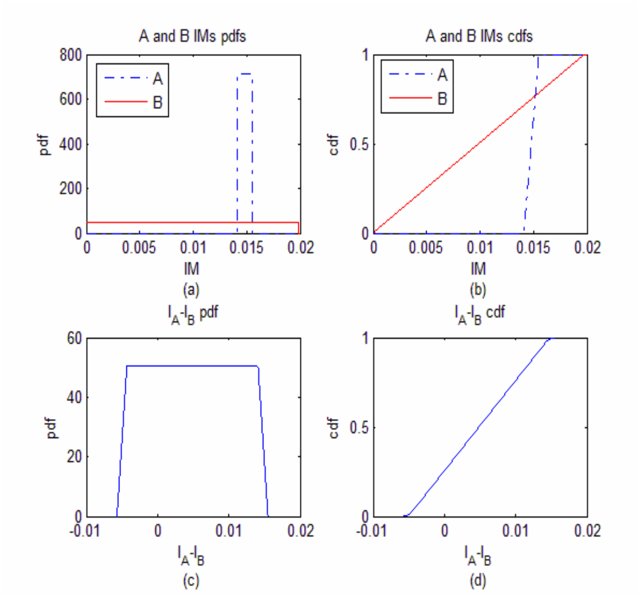

Figure 1. Probability density functions (pdf) and cumulative distribution functions (cdf) of the random variables $\mathbf{I_A}$, $\mathbf{I_B}$ (a and b) and $\mathbf{I_A}$-$\mathbf{I_B}$ (c and d) in case of IMs with uniformly distributed uncertainties.

The drawback of comparing the values corresponding to specific quantiles lies in the loss of information about the distribution: with reference for example to Figure 1, the fact that the 95th quantile of $\mathbf{I_A}$ (0.015) is lower than the 95th quantile of $\mathbf{I_B}$ (0.017) only means that the point value which $\mathbf{I_A}$ is lower than with probability of 0.95 is lower than the analogous point value for $\mathbf{I_B}$. The full information on the actual difference between the distributions of $\mathbf{I_A}$ and $\mathbf{I_B}$ does not play any role.

An obvious way to give due account to the difference between the distributions is to consider the random variable (rv) $\mathbf{I_A}$-$\mathbf{I_B}$ whose pdf and cdf are shown in Figures 1c and 1d, respectively. The details on their analytical expressions are given in Appendix 1. In order to establish if component A is more important than B, one can consider the probability 1-F_{AB}(0) that $\mathbf{I_A}$ is greater than $\mathbf{I_B}$; for example, in the present case r_{AB}=P($\mathbf{I_A}$>$\mathbf{I_B}$)=1-F_{AB}(0)=0.81, which means that with high probability component A is more important than B.

According to the above procedure for comparing the importance of two components A and B, it is necessary to fix a threshold T on the r_{AB} value: if 1- F_{AB}(0) is larger than T, then A is more important than B, otherwise no conclusion can be given. For example one may take T=0.5 or T=0.7; the lower the threshold the higher the risk associated with the decision. On the other hand, the choice of a crisp threshold as a probabilistic exceedance measure has some intrinsic limitations summarized in the following points:

- r_{AB} could fall just on T: in this case given the inevitable approximations and uncertainties related to the estimation of the IMs distributions, no robust conclusion can be given on the components importance (see Section below).
- Considering three components, A, B and C whose IMs are such that the cdf values in 0 of the IM differences fall very close to T, it could happen that $\mathbf{I_A}$>$\mathbf{I_B}$, $\mathbf{I_B}$>$\mathbf{I_C}$ and $\mathbf{I_C}$>$\mathbf{I_A}$.

These limitations can partially be overcome by taking not a crisp value of T but a range [T_l,T_u] (for example 0.4-0.6 or 0.3-0.7). Given two components A and B and the difference $\mathbf{I_A}$-$\mathbf{I_B}$:

- If r_{AB}> T_u, then A is more important than B.
- If r_{AB}< T_l, then B is more important than A.
- If T_l <r_{AB}< T_u, then A is equally important to B. In this case, different kinds of additional constraints/targets can guide the ranking order (costs, times, impacts on public opinion etc).

It may be of interest to relate the results provided by the probabilistic exceedance measure r_{AB}=P($\mathbf{I_A}$>$\mathbf{I_B}$) to the standard deviations of the IMs distributions, $\sigma_{\mathbf{IA}}$ and $\sigma_{\mathbf{IB}}$. *Figure 2* shows the variation of r_{AB} for increasing values of the standard deviations $\sigma_{\mathbf{IA}}$ and $\sigma_{\mathbf{IB}}$, keeping fixed the mean values of $\mathbf{I_A}$ and $\mathbf{I_B}$ and the ratio k=$\sigma_{\mathbf{IA}}$/$\sigma_{\mathbf{IB}}$. In the extreme case of no uncertainties on the knowledge of $\mathbf{I_A}$ and $\mathbf{I_B}$ ($\sigma_{\mathbf{IA}}$=0 and $\sigma_{\mathbf{IB}}$=0), component A is more important than B and thus r_{AB}=1. Increasing the standard deviations, r_{AB}=1 holds as long as the pdfs of $\mathbf{I_A}$ and $\mathbf{I_B}$ do not overlap, i.e. $\mathbf{I_A}$ and $\mathbf{I_B}$ are uncertain quantities but it is not uncertain that $\mathbf{I_A}$>$\mathbf{I_B}$. Finally, as the overlapping between pdfs increases r_{AB}

decreases. From these considerations one can argue that uncertainties on the IMs can affect the rank order and the reduction of the uncertainties should be, in certain cases, considered for decreasing the risk associated with the decision.

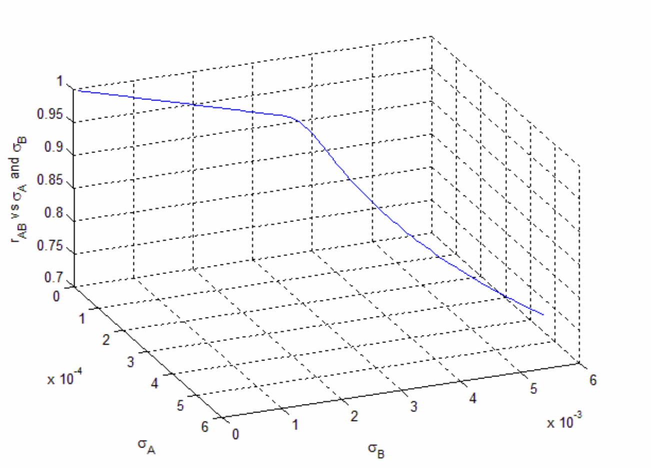

Figure 2. r_{AB} vs $\sigma_{\mathbf{IA}}$ and $\sigma_{\mathbf{IB}}$, keeping $\sigma_{\mathbf{IA}}/\sigma_{\mathbf{IB}}$, E[$\mathbf{I_A}$] and E[$\mathbf{I_B}$] constant.

3. An empirical procedure to rank component importance.

In the previous Section, a probabilistic measure of exceedance between two rvs has been utilized to compare components importance measures in presence of uncertainties. To extend the method to large numbers of components, a procedure for successive ranking must be introduced to avoid the combinatorial explosion of pairwise comparisons. The method proposed in this paper is an application of one of the most common sorting algorithms, Quicksort [5], which proceeds by choosing an element, called a pivot, and moving all smaller elements before the pivot and all larger elements after it. In the present case the order relation between the elements is guided by the measure of exceedance introduced in Section 2. To illustrate the method, two case studies have been examined:

1. failure rates lognormally distributed for a three components system; this case allows us to test the ranking criterion;
2. failure rates lognormally distributed for a more complex system; this case study allows us to explain in details the ranking procedure, with its advantages and limitations.

For simplicity only the Birnbaum IM is considered, the reasoning remaining exactly the same for the other IMs.

3.1 Three components system

The system, sketched in *Figure 3*, is made up of a series of 2 nodes: the first is constituted by 2 components in parallel logic, the second by a single component. To each of these components, a crisp reliability value has been assigned so that the IM values reported in *Table 2*, third column, for components A and B are the same as those considered in Section 2. Different IM values of the three components are given in *Table 2*, columns 3-6.

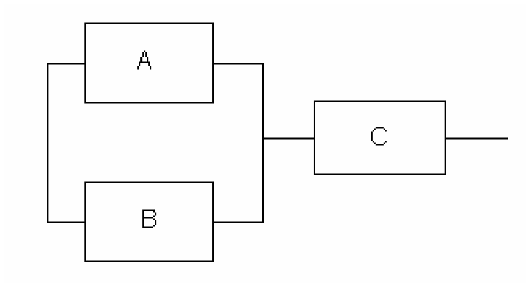

Figure 3. System Reliability Block Diagram.

Table 2. Components Reliability and Importance Measures.

	Reliability	Birnbaum	F-V	Criticality	RAW	RRW
A	0.988	0.0149	0.0012	0.0010	1.08	1.00
B	0.982	0.0099	0.0012	0.0010	1.05	1.00
C	0.825	0.9997	0.9989	0.9987	5.70	811

The components are assumed exponential, i.e. with constant failure rate λ_i, i=A, B, C. Uncertainties in the failure rates are described by lognormal distributions, (Figure 4, left), with parameters given in *Table 3*:

$$f_{\lambda_i}(\lambda_i) = \frac{e^{-\frac{[\ln(\lambda_i)-\mu_i]^2}{2\sigma_i^2}}}{\lambda_i \cdot \sigma_i \cdot \sqrt{2\pi}} \qquad (1)$$

At each time instant *t* the reliability of component *i* is:

$$r_i(\lambda_i, t) = e^{-\lambda_i t} \qquad (2)$$

with pdf (for 0<λ_i<1) (*Figure 4*, right):

$$f_{r_i}(\lambda_i, t) = -\frac{e^{-\frac{\left[\ln\left(-\frac{\ln(\lambda_i)}{t}\right)-\mu_i\right]^2}{2\sigma_i^2}}}{\lambda_i \cdot \ln(\lambda_i) \cdot \sigma_i \cdot \sqrt{2\pi}} \qquad (3)$$

The parameters of the distributions of the failure rates have been chosen such that the mean values of the reliability at time *t*=1200 (in arbitrary units of time) are equal to the values in column 3 of *Table 2*. In *Figures*

5 a and b, the pdfs of the failure rates and reliabilities at time *t*=1200 are reported for all three components.

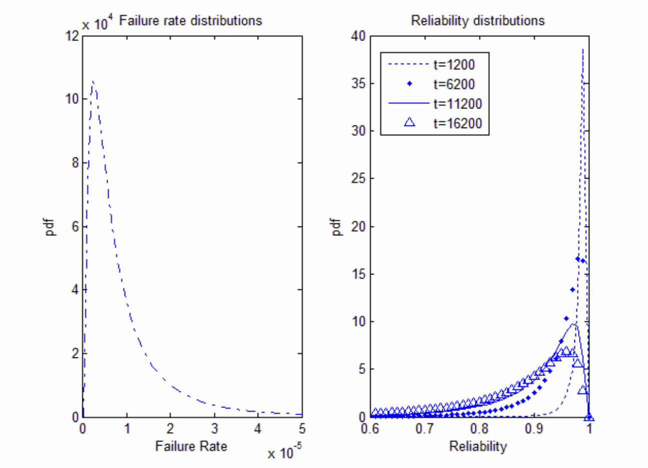

Figure 4. Lognormal distributions of the failure rate of component A (left) and corresponding pdfs of the reliability at different time instants (right).

Table 3. Parameters of the lognormal distributions of the components failure rates.

i	Mean	Variance
A	1.000e-005	1.649e-010
B	1.487e-005	1.076e-009
C	1.593e-004	1.694e-008

In spite of the simplicity of the considered system, finding the Birnbaum IM distributions by an analytical approach is impracticable. To overcome these difficulties Monte Carlo sampling has been applied. The resulting distributions are plotted in *Figures 5* c and d. It can be noted that the distribution of the IM of component C is displaced to larger values than A and B, which leaves no doubt that the most important component is C, as expected from the structure of the system. As for the ranking of A and B, one must compute the probability $P(\mathbf{I_A}>\mathbf{I_B})$.The result obtained by Monte Carlo sampling is r_{AB}=0.49997 which leads us to conclude that $\mathbf{I_A}=\mathbf{I_B}$. On the contrary, the Birnbaum IM values in *Table 2*, column 3, neglecting uncertainties, would lead to the conclusion that A is more important than B.

3.2 A more complex system

When the number of components in the system is large, the number of pairwise comparisons of their IMs increases dramatically. This calls for a systematic procedure of analysis to efficiently perform the importance ranking. Let us consider a system made up of a series of three nodes: the first is a 2-out-of-3 subsystem, the second consists of a single element and the third is a parallel system of two components. The Reliability Block Diagram (RBD) is reported in *Figure 6*. *Table 4* contains the data relative to the distribution of the components failure rates and reliabilities. The corresponding pdfs are plotted in *Figure 7*.

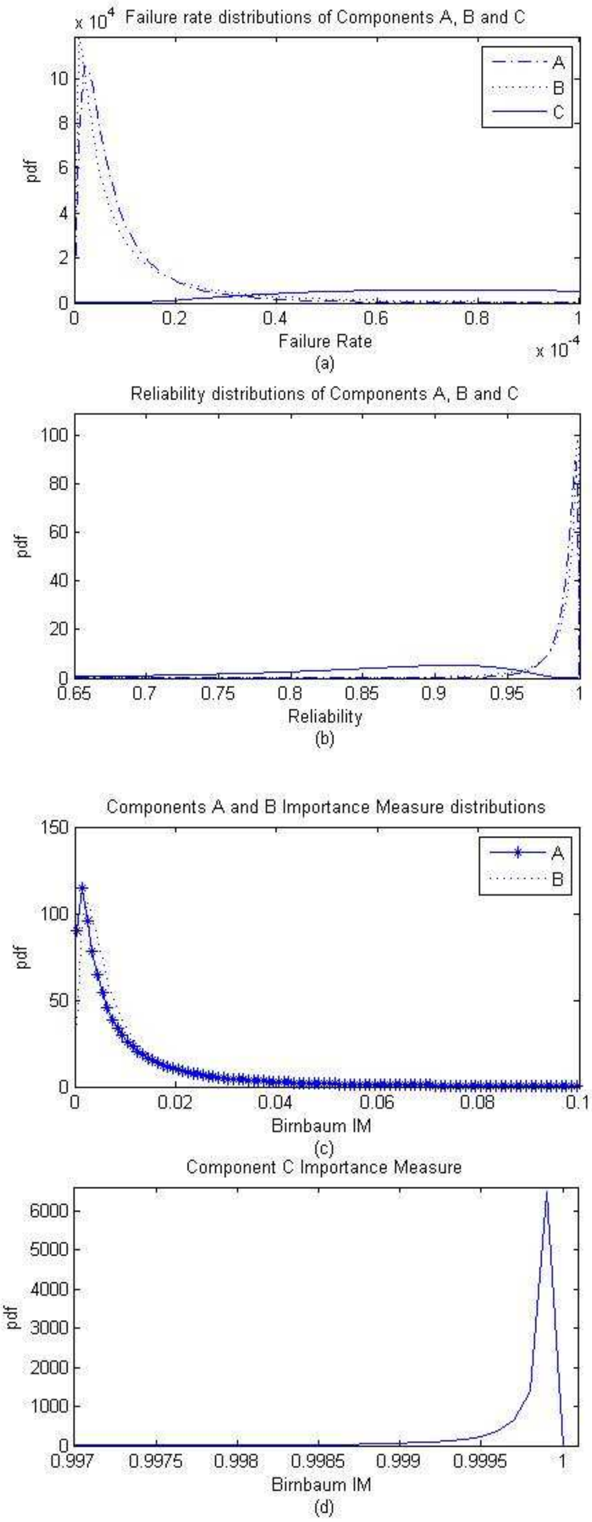

Figure 5. pdfs of the failure rate (a), reliability (b) and Birnbaum IM (c and d) of the three components.

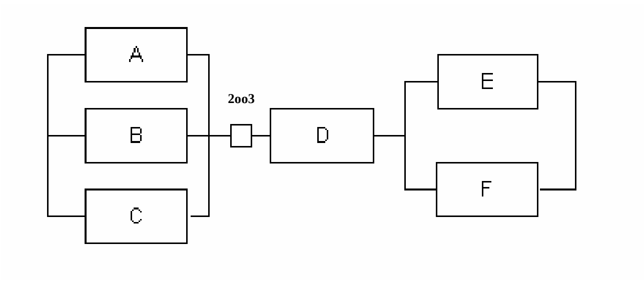

Figure 6. RBD of the system.

Table 4. Parameters of the lognormal distribution of the components failure rates.

i	Reliability	Mean	Variance	IM
A	0.98	8.37e-5	1.68e-8	0.056
B	0.97	1.26e-4	7.1e-10	0.048
C	0.96	1.69e-4	2.5e-8	0.040
D	0.825	7.97e-4	2.94e-6	0.99
E	0.988	5e-5	1.99e-8	0.015
F	0.982	7.52e-5	9.88e-9	0.009

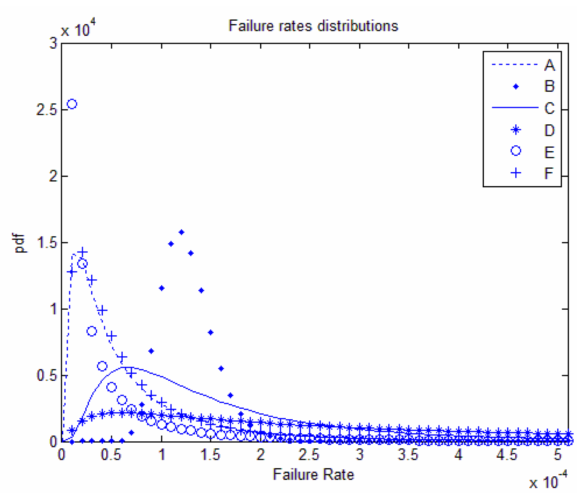

Figure 7. Failure Rate distributions.

The ranking procedure proposed in this Section is an application of one of the most famous sorting algorithms, Quicksort [5], to the probabilistic measure of exceedance $r_{ij}=P(\mathbf{I_i}>\mathbf{I_j})$ introduced in the previous Sections. Quicksort is a divide and conquer algorithm which relies on a partition operation: to partition an array, one chooses an element, called a pivot, moves all smaller elements before the pivot and all larger elements after it. In the iterative procedure, one then recursively sorts internally the sublists of smaller and larger elements. In the case of interest here, starting from the components rankings based on the reliability mean values, one chooses as pivot element the most important component, i.e. the one with larger importance measure $\mathbf{I_p}$ calculated based on the mean reliabilities. The $r_{pj}=P(\mathbf{I_p}>\mathbf{I_j})$ is calculated for each $j\neq p$ and the pre-defined threshold value T states the relation order between p and j, with respect to their $F_{pj}(0)$ cdf values.

Table 5. Column 1 reports the ranking obtained without considering uncertainties in the IMs; columns 2-6 the probability that the component in the row exceeds the component in the column; the last column shows the ranking obtained by the proposed method.

Mean ranking	A	B	C	E	F	Final rank
D	1	1	1	1	1	D
A		0.812	0.768	0.948	0.964	A
C		0.512		0.907	0.945	C,B
B				0.887	0.935	C,B
E					0.712	E
F						F

By proceeding this way, the components are ordered in function of their distance from the pivot element. When a swapping occurs between the ranks of two components, the check on the exceedance measure is repeated for all the components downstream the sublist. Doing so, it may happen that $R_i>R_j$, $R_j=R_k$ and $R_k=R_i$: in this case, i, j and k are considered equally important. The results obtained by applying this procedure on the system in *Figure 6* are reported in *Table 5*.
As expected from the structure of the system, component D is the most important one, followed by the components in the 2-out-of-3 subsystem. In particular, component A is the most important of this subsystem, then, components B and C result with the same importance; on the contrary, referring to the mean ranking, component B results more important than component C. Finally, components E and F (in the parallel subsystem of *Figure 6*) are prioritized accordingly to the general rule that the Birnbaum IM of components in parallel systems decreases with decreasing reliability of the components, both neglecting (*Table 4*, column 7) and accounting for uncertainties (*Table 5*, column 7).
As a term of comparison, the procedure proposed in [4] has been applied for the computation of an alternative measure of exceedance r_{ij}^*. It is based on the following two steps:

1. Estimation, for each component, of the probability of occupying a specific order in the ranking (probability mass function). This is achieved by performing Monte Carlo sampling: at each trial the components are given a rank order $\boldsymbol{R}_\mathbf{i}$ i=A, B, …,F. *Figure 8* provides the distributions of the obtained rank orders.
2. Computation of the measure of exceedance r_{ij}^* defined as:

$$r_{ij}^{*}=P(R_i>R_j)=\sum_{R_i=1}^{n}p(R_i)\sum_{R_j=1}^{R_i}p(R_j) \qquad (4)$$

The results obtained when applying this procedure are reported in *Table 6*.

Table 6. Rank order obtained by applying the alternative exceedance measure proposed in [4].

rank	mean ranking	A	B	C	E	F	Final Ranking
1	D	1	1	1	1	1	D
2	A		0.895	0.888	0.981	0.987	A
3	B			0.649	0.935	0.961	B,C
4	C		0.665		0.944	0.966	B,C
5	E					0.871	E
6	F						F

For what concerns the final ranking, the two methods considered provide the same results. Notice, however, that if one considers a different range [T_l,T_u], for example [0.4,0.6] instead of [0.3,0.7] the method proposed in [4] leads to ranking component B more important than C whereas with the method here proposed, B and C have the same importance. Moreover, if one considers a very small range, for example [T_l,T_u]=[0.49,0.51], then the method in [4] determines the ranking A, B, C for the IMs whereas the proposed method establishes A, C, B.

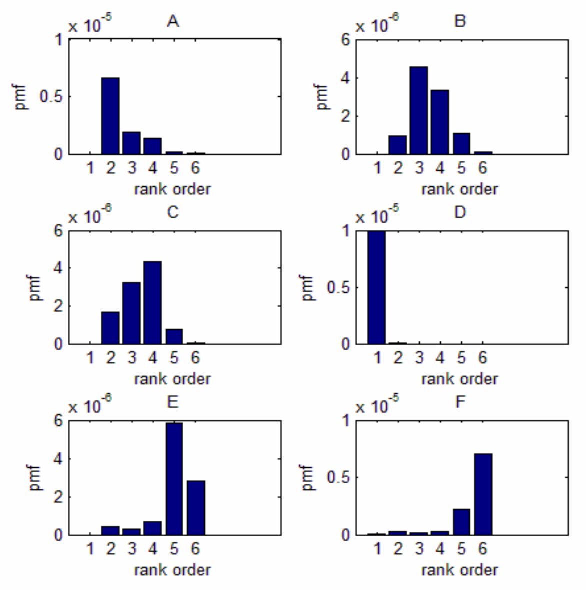

Figure 8. Probability mass functions of the rank order *R_i, i=A, …, F.*

The differences in the ranking of the two methods are caused by the different values of the exceedance measures between components B and C (*Tables 5* and *6*). In particular, the method proposed in this work results in r_{CB}=P($\mathbf{I_C}$>$\mathbf{I_B}$)=0.512, whereas the method in [4] results in r_{CB}^{*}=P($\mathbf{I_C}$>$\mathbf{I_B}$)=0.665. Notice that r_{CB} depends only on the importance measures of B and C whereas r_{CB}^{*} depends on the probability that a component occupies a specific order and thus also on the importance measures of the other components of the system. Moreover, another reason to prefer the procedure here proposed is that in the procedure proposed in [4], after each MC sampling the IMs of the components are used only to obtain a rank order, loosing the information on the actual numerical differences of the IMs. Finally, whereas for r_{ij} it holds that r_{ij} =1- r_{ji}, this property is not valid for r_{ij}^{*}, (for example, with reference to Table 6, r_{CB}^{*}=0.665 and $r_{BC}^{*}=0.649$). This means that the final ranking order may depend on the choice of the pivot in the Quicksort algorithm.

Conclusions and future works

In this work, a procedure is proposed for ranking system components in order of importance when in presence of uncertainties affecting the components reliabilities. The procedure is based on the definition of a probabilistic exceedance measure that permits to compare the importance of two components and can be summarized as follows:

1. Rank the components' importances according to their IMs computed by considering the expected values of their reliabilities, thus without considering uncertainties.
2. Define the range [T_l,T_u] of values of the probabilistic exceedance measure r_{pj}; for values r_{pj} in this range it is not possible to decide if $\mathbf{I_p}$>$\mathbf{I_j}$ or $\mathbf{I_p}$<$\mathbf{I_j}$ and this leads to consider components p and j as equally important.
3. Apply the Quicksort algorithm based on r_{pj}=P($\mathbf{I_p}$>$\mathbf{I_j}$):
   - 3.1. Put the components in the rank order found in step 1.
   - 3.2. Choose the first element of the list (sublist) as pivot element, p.
   - 3.3. For each j in the sublist compute the cdf, F_{pj}, of $\mathbf{I_p}$-$\mathbf{I_j}$ and evaluate r_{pj}=1- $F_{pj}(0)$:
     - If r_{pj} >T_u , then put j in the sublist of elements less important than p.
     - If r_{pj} <T_l , then put j in the sublist of elements more important than p.
     - If r_{pj} falls in [T_l,T_u], then p is equally important to j.

4. For each sublist go to 3.2 until no sublist with more than one element exists.

The application of the proposed procedure to two case studies has shown the importances of considering uncertainties in the computation of IMs: the ranking of the components' importance obtained neglecting the uncertainties affecting the component reliabilities can be different from the ranking obtained by considering uncertainties using the procedure proposed in this work.
Compared to another approach proposed in the literature, the procedure here presented seems to offer some advantages as, for example, the independence of the final rank from the choice of the pivot element in the Quicksort algorithm.

Acknowledgments

The work is funded by CESI RICERCA S.p.A. The views expressed in this paper are solely those of the authors.

References

[1] Birnbaum, Z. W. (1969). *On the importance of different components in a multi component system.* Multivariate analysis II, New York, Academic Press.
[2] Youngblood, R. W. (2001). Risk significance and safety significance. *Reliability Engineering and System Safety;* 73, 121-136.
[3] Apostolakis, G. E. (1990). The concept of probability in safety assessments of technological systems, *Science,* 250, 1359-1364.
[4] Modarres, M. (2006). *Risk Analysis in Engineering: Probabilistic Techniques, Tools and Trends*. CRC Press.
[5] Hoare, C. A. (1962). Quicksort. *Computer Journal*, 5, 10-15.
[6] Zio, E. (2006). *An introduction to the basics of reliability and risk analysis.* Singapore, World Scientific Publishing.
[7] Rausand, M. & Hoyland, A. (2004). *System Reliability Theory*. Wiley.
[8] Birolini, A. (2004). *Reliability Engineering.* Springer.
[9] Schneeweiss, W. G. (2001). *Reliability Modeling.* LiLoLe-Verlag.

Appendix 1

Given a generic uniform rv **x**≈U(a,b), its moment generating function (mgf) is given by:

$$\phi(s) = \frac{e^{sb} - e^{sa}}{s(b-a)}$$

The rv $\mathbf{r}=\mathbf{I_A}-\mathbf{I_B}$ is the convolution of two uniformly distributed random variables and in particular:

$\mathbf{I_A}$≈U(a_{IA},b_{IA});

$-\mathbf{I_B}$≈U($-b_{IB}$,$-a_{IB}$);

The mgf of **r** is given by:

$$\phi_r(s) = \frac{e^{sb_{I_A}} - e^{sa_{I_A}}}{s(b_{I_A} - a_{I_A})} \cdot \frac{e^{s(-a_{I_B})} - e^{s(-b_{I_B})}}{s(b_{I_B} - a_{I_B})}$$

$$\phi_r(s) = \frac{e^{s(b_{I_A}-a_{I_B})} - e^{s(a_{I_A}-a_{I_B})} - e^{s(b_{I_A}-b_{I_B})} + e^{s(a_{I_A}-b_{I_B})}}{s^2(b_{I_A} - a_{I_A}) \cdot (b_{I_B} - a_{I_B})}$$

For what concerns the inverse transformation, it could be noted that **r** mgf can be regarded as the algebraic sum of functions linearly increasing/decreasing with the same slope. So the pdf and cdf of **r** are given by:

$$f_r(r) = \begin{cases} \dfrac{r + b_{I_B} - a_{I_A}}{(b_{I_A} - a_{I_A}) \cdot (b_{I_B} - a_{I_B})} & a_{I_A} - b_{I_B} \le r \le b_{I_A} - b_{I_B} \\ \dfrac{1}{(b_{I_B} - a_{I_B})} & b_{I_A} - b_{I_B} \le r \le a_{I_A} - a_{I_B} \\ \dfrac{b_{I_A} - a_{I_B} - r}{(b_{I_A} - a_{I_A}) \cdot (b_{I_B} - a_{I_B})} & a_{I_A} - a_{I_B} \le r \le b_{I_A} - a_{I_B} \end{cases}$$

$$F_r(r) = \begin{cases} \dfrac{(b_{I_B} - a_{I_A} + r)^2}{2 \cdot (b_{I_A} - a_{I_A}) \cdot (b_{I_B} - a_{I_B})} & a_{I_A} - b_{I_B} \le r \le b_{I_A} - b_{I_B} \\ 1 - \dfrac{(a_{I_A} + b_{I_A} - 2 \cdot a_{I_B} - 2 \cdot r)}{2 \cdot (b_{I_B} - a_{I_B})} & b_{I_A} - b_{I_B} \le r \le a_{I_A} - a_{I_B} \\ 1 - \dfrac{\frac{(b_{I_A} - a_{I_B} - r)^2}{2}}{(b_{I_A} - a_{I_A}) \cdot (b_{I_B} - a_{I_B})} & a_{I_A} - a_{I_B} \le r \le b_{I_A} - a_{I_B} \end{cases}$$