Asociación Argentina
de Mecánica Computacional

Mecánica Computacional Vol XXXVI, págs. 723-739 (artículo completo)
José G. Etse, Bibiana M. Luccioni, Martín A. Pucheta, Mario A. Storti (Eds.)
San Miguel de Tucumán, 6-9 Noviembre 2018

HOW TO DEAL WITH UNCERTAINTY QUANTIFICATION AND PROPAGATION

Roberta Lima and Rubens Sampaio

Department of Mechanical Engineering, PUC-Rio, Rua Marquês de São Vicente, 225, Gávea, 22451-900 RJ, Brazil

Keywords: Uncertainty quantification, uncertainty propagation, measure of uncertainty, Shannon entropy, moments, envelope graph.

Abstract. Over the last years, stochastic models have called the attention of researchers. The number of publications in the subject has grown and the topic is being analyzed in different applications. Observing the new literature produced in the field, it is possible to verify that new expressions, as uncertainty quantification and uncertainty propagation, have emerged. These expressions become largely used and fashionable, however there is no consensus of their meanings. Each author arbitrates a meaning according to its own convenience. Sometimes the disarray is such that different meanings, some of them contradictory, can be found throughout the same work. The objective of this paper is to clarify the concepts. We define what is and what is not uncertainty quantification and propagation. We also show, with simple examples, that several strategies found in literature called strategies to compute uncertainty quantification and propagation are not. They can lead to errors and misleadingness. The examples were chosen to be as simple as possible in order to highlight different problems that can arise when one uses these strategies.

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar

1 INTRODUCTION

The literature dealing with stochastic models has grown over the last years. The number of papers, books and courses in the subject increased Ghanem et al. (2017); Le Maître and Knio (2010); Smith (2014); Soize (2017, 2012); Sullivan (2015) and Observing the new literature produced in the field, it is possible to verify that new expressions, as uncertainty quantification and uncertainty propagation, have emerged. Since these expressions do not appear in classical books and papers and nowadays they are widely used, one may ask: what are their meaning? Reading the literature that use them, to answer this question is not a simple task. Different papers employ them with different meanings. There is no consensus in the literature. Each author arbitrates a meaning according to its own convenience. Sometimes the disarray is such that different meanings, some of them contradictory, can be found throughout the same work.

The objective of this paper is to clarify the concepts. We define what is and what is not uncertainty quantification and propagation Besides of this, we show that several strategies found in the literature called strategies compute uncertainty quantification and propagation are not.

This paper is organized as follows. Section 2 presents the definition uncertainty quantification and Section 3 presents some problems that may arise with several strategies found in the literature that are used to quantify uncertainty. The problems are classified into 3 types: dimension, meaning and incompatibility. Some examples of incompatibility are shown in Section 4. Section 5 presents the definition of uncertainty propagation and, it is shown how this definition is associated with random processes. In Section 6 three strategies found in the literature to measure uncertainty propagation with a set of statistics are discussed. To illustrate the problems that can arise with the use of these strategies, some simple examples are given in Section 7.

2 DEFINITION OF UNCERTAINTY QUANTIFICATION

Uncertainty is, certainly, described by the cumulative distribution function (CDF). Since all random variables and random vectors have a CDF, one associates an uncertainty to each of them. Using the CDF, one describes the three main cases and their combinations: when there is an absolutely continuous, with respect to the Lebesgue measure, probability density function, the discrete case, and the singular case. That is, of course, the reason why one does not see any mention to uncertainty quantification in classical books as Feller (1957); Chung (1974). The authors saw no reason to call a CDF by another name. CDF is uncertainty. The prescription of a CDF is its quantification.

However, one has to acknowledge that to use a CDF to describe uncertainty is clumsy. One feels there must be a simpler way. Why not use some small set of statistics to reduce a CDF to a simpler measure, easier to grasp? This seems a great idea and, indeed, one finds it in the literature. It is common to find papers using statistics as measures. In the case of random variables, the statistics most used are mean, variance, coefficient of variation and Shannon entropy. The idea is to associate numbers to the CDF, as show in Fig. 1.

We focus the discussion on three main cases:

1. to use the Shannon entropy;

2. to use mean and standard deviation to construct an envelope and with them to make a nice graph;

3. to use mean and coefficient and variation;

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar

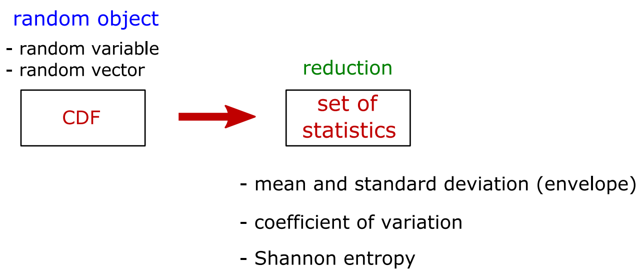


Figure 1: Reductions: to replace the CDF for a small set of statistics.

In the probabilistic context, the Shannon entropy, S, Jaynes (1957); Shannon (1948) of a random variable is viewed as a measure of the information carried by the associated probabilistic distribution. Sometimes, the Shannon entropy is used as a synonym of uncertainty. In this paper uncertainty is the CDF, entropy is a statistics computed from a CDF. It reflects some properties of the CDF, but not all. In the case of discrete random variables, S is defined using the mass function. It is dimensionless. In the case of absolutely continuous random variables using the probability density function. It is not dimensionless.

For some probabilities, one can associate a mean μ, variance σ^2, and coefficient of variation $\delta = \sigma/\mu$ (ratio between the standard deviation and mean). The mean of a probability mass, or distribution, is the best approximation of it by a number, and the absolute error of this approximation, in the mean square sense, is the variance. The coefficient of variation is a measure of the relative error.

It is common to find papers using statistics as measures of uncertainty, as we can see in Chen et al. (2010); Khosravi and Nahavandi (2014); Nordström and Wahlsten (2015); Motra et al. (2016); Zidek and Eeden (2003); Conrad (2016). Please note that statements like *given a random variable with fixed mean, when the variance grows, the level of uncertainties also grows* or *for two random variables with different means and variances, the random variable with the higher coefficient of variation is always the more uncertain* are meaningless.

The reductions - to replace the CDF for a small set of statistics - may indeed work in some cases. But they do not always work and, moreover, the different measures they define may not be compatible. That is, the ordering of uncertainty, in the measure defined by the statistics, may vary depending on what set of statistics one chooses. So, the great idea does not work so far, but it is happily used in the literature.

3 STRATEGIES FOUND IN THE LITERATURE TO QUANTIFY UNCERTAINTY

As said in the introduction, to use a CDF to describe uncertainty is clumsy. One of the problems that arise with this description is that the comparison of CDF to see which is more uncertain is not evident. The idea found in the literature to solve this dilemma is to use some small set of statistics to reduce a CDF to a simpler measure, easier to grasp. The most used statistics are variance, coefficient of variation and Shannon entropy.

The reductions - to replace the CDF for a small set of statistics - may indeed work in some cases. However, these reductions present problems in relation to the dimension, meaning, and incompatibility of statistics.

3.1 Dimension of statistics

It would be useful if the same set of statistics could be used to all random objects, random variables (discrete and continuous) and random vectors for instance. However, the statistics of

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar

different objects have different dimension, in a way that it is not possible to compare them.

- For random variables, the mean, variance, coefficient of variation, and entropy are scalars. Beside this, the entropy of discrete random variables is always a non-negative value and the entropy of continuous random variables can assume values in $\mathbb{R}$.
- For random vectors, the mean and variance are vectors. The covariance and correlation are matrices and the entropy is a scalar. Observe that for random vectors the coefficient of variation is not defined in the literature.

Figure 2 outlines the dimensions. Another problem is that there are random variables that do not have any moment, for example the random variables with Cauchy density function.

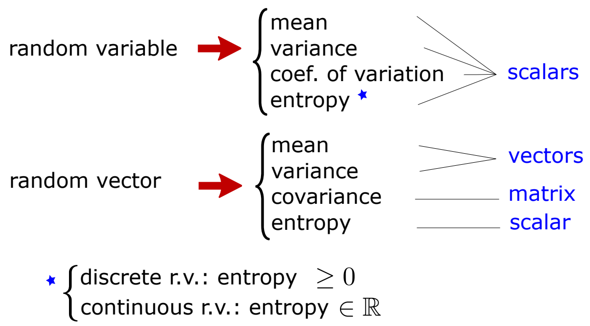


Figure 2: Dimension of statistics.

3.2 Meaning of statistics

The statistics usually used to reduce the CDF are variance, coefficient of variation and Shannon entropy. However, these statistics already have their own meaning. The mean of a probability mass, or distribution, is the best approximation of it by a number, and the norm of the quadratic error of this approximation, in the mean square sense, is the variance. The coefficient of variation is a measure of the relative error. The Shannon entropy of a random variable is viewed as a measure of the information carried by the associated probabilistic distribution.

3.3 Incompatibility

Another problem that appears in the use of a set of statistics as measure of uncertainty is that the different measures may not be compatible. That is, the ordering of uncertainty may vary depending what set one chooses. In the next section, some examples show the contradictions.

Beside this, there is an incompatibility in relation to the entropy of discrete random variables and the entropy of continuous random variables. Observe that while the entropy of discrete random variables is always a non-negative value, the entropy of continuous random variables can assume values in $\mathbb{R}$. Therefore, the lower value of entropy of a discrete random variable is 0 and of a continuous random variable is $-\infty$.

For example: if X is uniformly distributed in $[0,\ 2^n]$, then the entropy $S_X = -\int_0^{2^n} \frac{1}{2^n} \ln\left(\frac{1}{2^n}\right) dx = \ln 2^n$. Also, if Y is uniformly distributed in $[0,\ 2^{-n}]$, then $S_Y = \ln 2^{-n}$. Observe that as the support of the distribution increases, the entropy increases. In the limit, it goes to ∞. As the support decreases, the entropy decreases and, in the limit, it goes to $-\infty$. This case where the

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar

support tends to zero could be seen as an limit case, in which the random variable assumes only one value. This is certainty, deterministic. Therefore, for continuous random variables, if entropy is considered a measure of uncertainty, the lower measure (corresponding to the certainty) is $-\infty$. However, the limit obtained for discrete random variables is different. If we consider a discrete random variable which assumes only one value, we get that the entropy of certainty is 0. The conclusion is that entropy could not be used as measure of uncertainty of random variables. That is no consensus of what is the entropy of certainty.

4 EXAMPLES OF INCOMPATIBILITY IN THE STRATEGIES FOUND IN THE LITERATURE TO QUANTIFY UNCERTAINTY

4.1 Bimodal density function

Consider a family of continuous random variables X_d, parameterized by d, with a bimodal density function p_d symmetrically distributed around its mean μ. The function p_d, sketched in Fig. 3, is

$$p_d(x) = \begin{cases} 1, & x \in [\mu - (d/2) - (1/2)\,, \mu - (d/2)]\,, \\ 1, & x \in [\mu + (d/2)\,, \mu + (d/2) + (1/2)]\,, \\ 0, & \text{in all others cases.} \end{cases} \tag{1}$$

The variance of this family of random variables X_d is

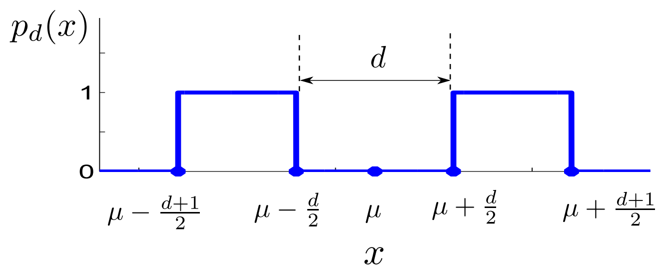


Figure 3: Family of bimodal density functions p_d.

$$\sigma_d^2 = E[(X_d - \mu)^2] = \frac{1}{12}(3d^2 + 3d + 1)\,. \tag{2}$$

The entropy is

$$S_d = -\int_{-\infty}^{\infty} p_d(x) \ln p_d(x)\, dx = 0\,. \tag{3}$$

Observing Eqs. (15), and (16), we verify that for a given value of μ, as d (distance between peaks of the bimodal density function) increases, the variance, σ_d^2, and the coefficient of variation, $\delta_d = \sigma_d/\mu$ (for $\mu \neq 0$), increase. However, the entropy S_d remains constant and equal to zero.

This example of the bimodal family shows quite well that entropy, mean, variance, and coefficient of variation are different things. One can vary μ, σ^2, $\delta = \sigma/\mu$ independently with fixed entropy.

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar

4.2 Gaussian density function

The next example deals with a continuous random variable. Consider a family of continuous random variables X_μ, parametrized by μ, with a Gaussian density function p_μ

$$p(x)_\mu = \frac{1}{\sqrt{2\sigma^2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}\ , \tag{4}$$

where μ is the mean and σ^2 the variance. The entropy of the Gaussian density function is

$$S_\mu = -\int_{-\infty}^{\infty} p(x)\ln p(x)\, dx = \frac{1}{2}\ln\left(2\sigma^2\pi\, e\,\right). \tag{5}$$

Observing Eq. (5), we verify that the mean μ does not enter the final formula of the entropy. This means that all Gaussian functions with a common variance σ^2 have the same entropy. A translation changes μ and $\delta = \sigma/\mu$, but does not change the value of S_μ. Increasing the mean, δ decreases, and decreasing the mean, δ increases.

4.3 Gamma density function

The next example also shows that it is possible that an increase of the variance corresponds to a decrease of entropy. Consider a family of continuous random variables X with a Gamma density function p

$$p(x) = \frac{1}{\Gamma(k)\theta^k} x^{k\,-\,1} e^{-\frac{x}{\theta}}\ . \tag{6}$$

written as function of the parameters $k > 0$ and $\theta > 0$, a shape parameter and a scale parameter, respectively. The mean of X is $\mu = k\theta$ and the variance $\sigma^2 = k\theta^2$. The entropy is Khodabin and Ahmadabadi (2010)

$$S = -\int_{-\infty}^{\infty} p(x)\ln p(x)\, dx = k + \ln\theta + \ln[\Gamma(k)] + (1\ -\ k)\psi(k). \tag{7}$$

where ψ is the digamma function. In Fig. 4 it is shown the graph of the entropy S as function of the variance σ^2 for different values of the mean μ of X. We verify that for the smaller value of the mean $\mu = 1.0$, the value of S decreases as the variance increases. When $\mu = 3.0$, an interesting behavior can be observed. For values of variance lower then 9.0, S increases as the variance increases. However, for values of variance above 9.0, S decreases as the variance increases. This means that entropy and σ^2 may not vary in the same sense. Thus, if the Shannon entropy S is considered a measure of the level of uncertainties, the variance can not be taken as a measure of the level of uncertainties either.

5 DEFINITION OF UNCERTAINTY PROPAGATION

As CDF is uncertainty, uncertainty propagation is the variation of the CDF with a parameter (discrete or continuous), i.e., the dynamics of the CDF, as sketched in Fig. 5. Remark that this problem has already been studied by Einstein years ago in the context of Brownian motion. Einstein formulated this problem through a diffusion equation Einstein (2018), an equation that describes how is the dynamics of the probability distribution function, PDF, of the position of one particle. Integrating it, one can describe the variation of the CDF.

The literature around 1905 does not not mention the expression uncertainty propagation. Despite dealing with the problem, the authors of the time did not call it by uncertainty propagation.

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar

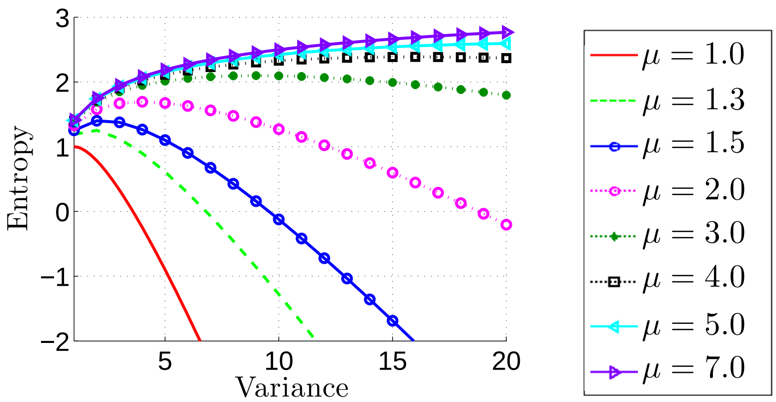


Figure 4: Entropy as function of the variance of a continuous random variable with a Gamma density function.

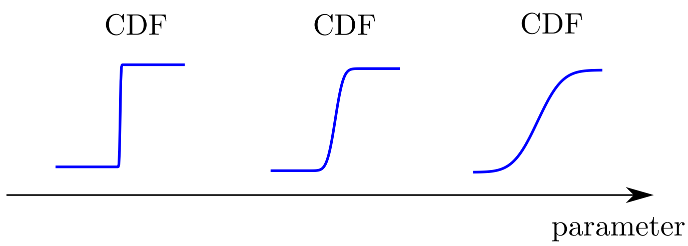


Figure 5: Uncertainty propagation: the dynamics of the CDF.

Only recently the expression appeared and spread in papers, books and courses. The expression is new, but what it represents is not!

Nowadays, uncertainty propagation is a frequent term in the literature. Several papers claim that they compute uncertainty propagation. However, what it is actually done, in the majority of them, is anything but it. Instead of to compute the variation of the CDF with a parameter, they compute the variation of a set of statistics with a parameter. In this paper we focus the discussion on three strategies found in the literature called "uncertainty propagation":

1. to compute the variation of mean and standard deviation and then to use them to construct an envelope graph;
2. to compute the variation of coefficient of variation;
3. to compute the variation of Shannon entropy.

To replace the CDF by a small set of statistics seems to be a powerful tool to determine uncertainty propagation. Instead of dealing with the CDF, which is a function, one deals with a reduction of it, statistics. With the reduction, it would be possible to determine what is happening with the uncertainty and, some questions as "What is happening with the uncertainty?, Is it growing or decreasing?" could be answered. However what apparently is a powerful tool can lead to errors and misleadingness. One possible misleadingness is that that different strategies may not be compatible, moreover, they can be contradictory. While a set statistics suggests that the uncertainty is growing with a parameter, another set suggests that the uncertainty is decreasing. Each set of statistics may provide different information about what is happening with the uncertainty.

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar

Since one of the objectives of this paper is to discuss the three strategies found in the literature based on reductions and, to show with simple examples how they can mislead. The examples were chosen to be as simple as possible. Examining them one sees immediately the inadequacy of the reductions.

6 STRATEGIES FOUND IN THE LITERATURE TO COMPUTE UNCERTAINTY PROPAGATION

In this Section, three strategies, frequently called in the literature as uncertainty propagation are presented. The first one is based on the first two moments of a probability distribution: mean and variance (square of the standard deviation). As the mean value is interpreted as a nominal value and the standard deviation is the error Souza de Cursi and Sampaio (2015), we will call this strategy as absolute error analysis. This first strategy consists in to compute the variation of mean and standard deviation with a parameter to construct a region showing three curves. Two bounding curves showing the mean $\pm$ standard deviation and the third curve, between the bounding two, that shows the mean value, seen as a nominal value. This procedure constructs a graph typically called envelope. A second strategy is based on coefficient of variation (ratio between standard deviation and mean), seen as a measure of relative error. The idea is to compute the variation of the coefficient of variation. Remark that these two strategies may not always be used. In some cases, neither envelope nor coefficient of variation exist, as in the random process $\mathcal{X}(t) = t\,C$, with $t \in \mathbb{R}$, where C is a random variable with Cauchy distribution. Finally, the third strategy usually found in literature is to compute the variation of entropy. Although these three strategies use different sets of statistics, it is a common mistake to believe that they provide the same information. Next we present some simple examples where one verifies that they may not be compatible, moreover, they may be contradictory.

7 EXAMPLES OF INCOMPATIBILITY IN THE STRATEGIES FOUND IN THE LITERATURE TO COMPUTE UNCERTAINTY PROPAGATION

.

7.1 Example 1:

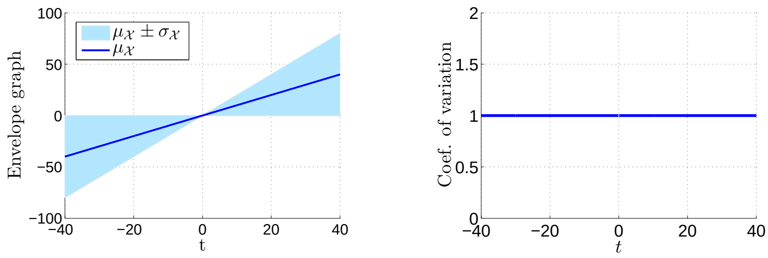
Given a random variable Y with Gaussian distribution (with mean $\mu_Y \neq 0$ and variance σ_Y^2), consider the random process $\mathcal{X}(t) = t\,Y$, with $t \in \mathbb{R}$. For a specific $t \in \mathbb{R}, t \neq 0$, $\mathcal{X}(t)$ is a random variable with Gaussian distribution with mean $\mu_{\mathcal{X}}(t) = t\,\mu_Y$, variance $\sigma_{\mathcal{X}}^2(t) = t^2\,\sigma_Y^2$, coefficient of variation $\delta_{\mathcal{X}}(t) = \dfrac{\sigma_{\mathcal{X}}(t)}{\mu_{\mathcal{X}}(t)} = \delta_Y$ and, entropy $\eta_{\mathcal{X}}(t) = \dfrac{1}{2}\ln\left(2\pi e \sigma_{\mathcal{X}}^2(t)\right) = \dfrac{1}{2}\ln\left(2\pi e t^2\,\sigma_Y^2\right)$.

Figures. 6(a), 6(b) and 7(a) show an envelope graph, coefficient of variation and entropy of $\mathcal{X}$ as function of the parameter t for $\mu_Y = 1.0$ and $\sigma_Y^2 = 1.0$. Remark that the envelope suggests that the "uncertainty" decreases when t approaches zero. However, the coefficient of variation suggests that the "uncertainty" is constant. The entropy gives a third different behaviour. As t approaches zero from the positive or negative side, the entropy goes to $-\infty$. This simple example shows the contradictions of the strategies. To compute the uncertainty propagation, it is necessary to compute the variation of the CDF of $\mathcal{X}(t)$ with t, given by

$$P_{\mathcal{X}(t)}(x) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x - \mu_{\mathcal{X}(t)}}{\sigma_{\mathcal{X}(t)}\sqrt{2}}\right)\right] = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x - t\,\mu_Y}{t\,\sigma_Y\sqrt{2}}\right)\right]. \tag{8}$$

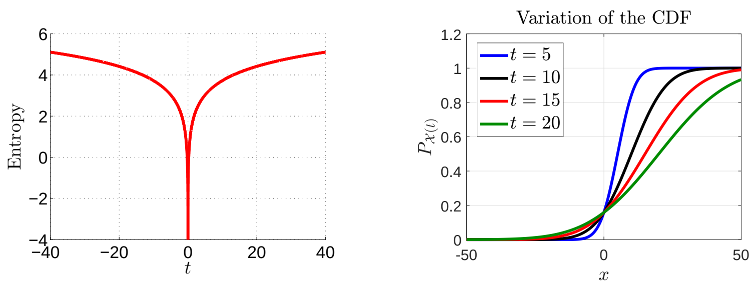
Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar

Some examples are given Fig. 7(b).



(a) (b)

Figure 6: (**a**) Envelope graph and (**b**) coefficient of variation of $\mathcal{X}$ as function of t.



(a) (b)

Figure 7: (**a**) Entropy of $\mathcal{X}$ as function of t and (**b**) cumulative distribution function (CDF) of $\mathcal{X}(t)$ for different values of the parameter t.

7.2 Example 2:

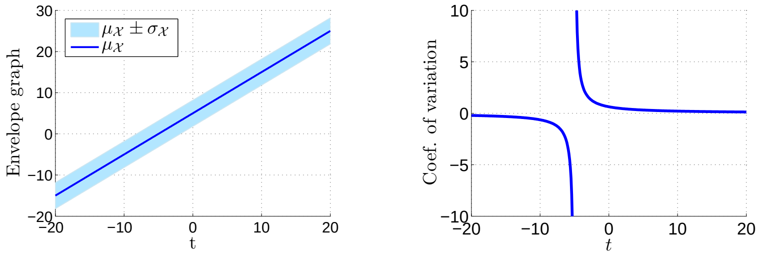
Given a random variable Y with Gaussian distribution (with mean μ_Y and variance σ_Y^2), consider the random process $\mathcal{X}(t) = t + Y$, with $t \in \mathbb{R}$. For a specific $t \in \mathbb{R}$, $\mathcal{X}(t)$ is a random variable with Gaussian distribution with mean $\mu_{\mathcal{X}}(t) = t + \mu_Y$, variance $\sigma_{\mathcal{X}}^2(t) = \sigma_Y^2$, coefficient of variation $\delta_{\mathcal{X}}(t) = \dfrac{\sigma_{\mathcal{X}}(t)}{\mu_{\mathcal{X}}(t)} = \dfrac{\sigma_Y}{\mu_Y + t}$ and, entropy $\eta_{\mathcal{X}}(t) = \dfrac{1}{2}\ln\left(2\pi e \sigma_{\mathcal{X}}^2(t)\right) = \dfrac{1}{2}\ln\left(2\pi e t^2\, \sigma_Y^2\right)$.

Figures. 8(a), 8(b) and 9(a) show an envelope graph of $\mathcal{X}$, coefficient of variation and entropy as function of the parameter t for $\mu_Y = 5.0$ and $\sigma_Y^2 = 10.0$. Each graph exhibits a different behavior. While the envelop suggests a linear growth with t, the coefficient of variation is smaller than zero when $t < -\mu_Y$ and it is bigger then zero when $t > -\mu_Y$. The entropy, similar to the previous example, suggests that the "uncertainty" is constant. The CDF of $\mathcal{X}(t)$ with t is given by

$$P_{\mathcal{X}(t)}(x) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x - \mu_{\mathcal{X}(t)}}{\sigma_{\mathcal{X}(t)}\sqrt{2}}\right)\right] = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x - (t + \mu_Y)}{\sigma_Y\sqrt{2}}\right)\right]. \quad (9)$$

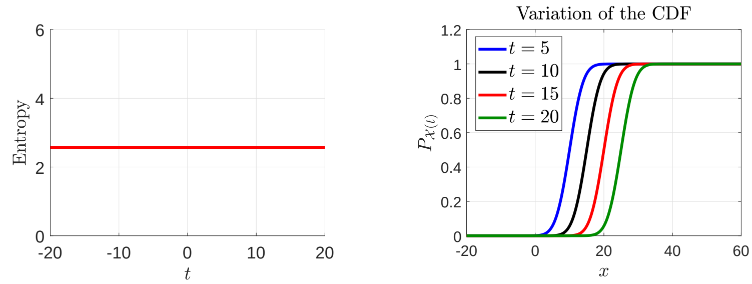
Some examples are given Fig. 9(b).

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar



(a) (b)

Figure 8: (**a**) Envelope graph and (**b**) coefficient of variation of $\mathcal{X}$ as function of t.



(a) (b)

Figure 9: (**a**) Entropy of $\mathcal{X}$ as function of t and (**b**) cumulative distribution function (CDF) of $\mathcal{X}(t)$ for different values of the parameter t.

7.3 Example 3:

Consider the random process $\mathcal{X}(t) = t\,C$, with $t \in \mathbb{R}$, where C is a random variable with Cauchy distribution:

$$p_C(x) = \frac{1}{\pi\gamma}\left[\frac{\gamma^2}{(x-x_0)^2+\gamma^2}\right], \tag{10}$$

where x_0 is a location parameter, specifying the location of the peak of the distribution, and γ is a scale parameter. When $x_0 = 0$ and $\gamma = 1$, the cumulative distribution function, the characteristic function and entropy of C are respectively:

$$P_C(x) = \frac{1}{\pi}\arctan\left(\frac{x}{\gamma}\right) + \frac{1}{2}, \tag{11}$$

$$\phi_C(a) = E[e^{iaX}] = e^{-|a|}, \qquad \forall a \in \mathbb{R}, \tag{12}$$

$$\eta_C = \ln(4\pi\gamma). \tag{13}$$

For a specific $t \in \mathbb{R}$, $\mathcal{X}(t)$ is a random variable with Cauchy distribution, thus it does not have any moments. In this case, it is not possible to construct neither envelop nor coefficient of variation graphs.

7.4 Example 4: free vibration

Consider a simple mass-spring oscillator moving on a horizontal surface without friction, as shown in Fig. 10.

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar

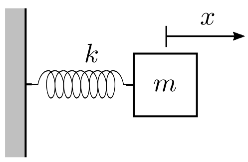


Figure 10: Free mass-spring oscillator.

Its equation of motion is

$$m\,\ddot{x}(t) + k\,x(t) = 0\,. \tag{14}$$

where m is the mass, k is the spring stiffness, t is time and x is the position of the mass. Considering that $m = 1.0$ kg and the initial conditions are $x(0) = 1$ m and $\dot{x}(0) = 0$ m/s, the solution of Eq. 14 is

$$x(t) = \cos\,(\sqrt{k}t)\,. \tag{15}$$

Consider that the spring stiffness is uncertain and modeled as a discrete random variable K with Bernoulli distribution. It takes the values 1 N/m and $\pi^2/4$ N/m, each one with probability $1/2$. The mass function of K is

$$\begin{aligned} p(K = 1) = 1/2\,, \\ p(K = \pi^2/4) = 1/2\,. \end{aligned} \tag{16}$$

As it was assumed that the spring stiffness is uncertain, the response of the system is a random process, $\mathfrak{X}$ Grimmett and Welsh (1986); Sampaio and Lima (2012). Since K has Bernoulli distribution, for each value of $t > 0$, $\mathfrak{X}(t)$ has also a Bernoulli distribution. With probability $1/2$, $K = 1$ N/m and then $\mathfrak{X}(t) = \cos\,(t)$. Also with probability $1/2$, $K = \pi^2/4$ N/m and then $\mathfrak{X}(t) = \cos\,(\frac{\pi}{2}t)$. Then the mean, variance and entropy are, respectively:

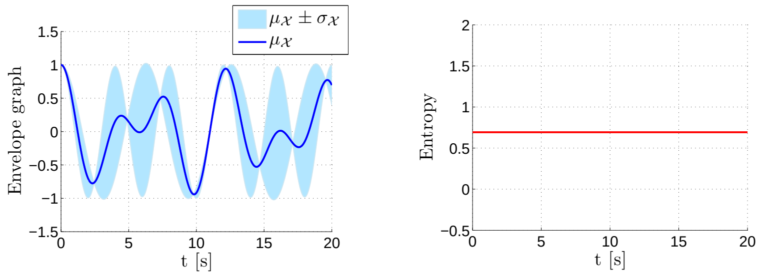
$$\mu_{\mathfrak{X}}(t) = E[\mathfrak{X}(t)] = \frac{1}{2}\cos\,(t) + \frac{1}{2}\cos\,(\frac{\pi}{2}t)\,, \tag{17}$$

$$\begin{aligned} \sigma^2_{\mathfrak{X}}(t) =& E[(\mathfrak{X}(t) - \mu_{\mathfrak{X}}(t))^2] = E[\mathfrak{X}^2(t)] - \mu^2_{\mathfrak{X}}(t) \\ =& \frac{1}{2}\left\{\cos^2\,(t) + \cos^2\,(\frac{\pi}{2}t)\right\} - \frac{1}{4}\cos^2\,(t) - \frac{1}{4}\cos^2\,(\frac{\pi}{2}t) - \frac{1}{2}\cos\,(t)\cos\,(\frac{\pi}{2}t) \\ =& \frac{1}{4}\left\{\cos\,(t) - \cos\,(\frac{\pi}{2}t)\right\}^2\,. \end{aligned} \tag{18}$$

$$\eta_{\mathfrak{X}}(t) = -\frac{1}{2}\ln\left(\frac{1}{2}\right) + \frac{1}{2}\ln\left(\frac{1}{2}\right) \approx 0.69 \quad \forall t > 0\,. \tag{19}$$

Comparing the envelope and entropy graph, shown in Figs. 11(a) and 11(b), one sees, immediately, that they suggest different behaviors.

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar



(a) (b)

Figure 11: (**a**) Envelope graph of the system response. (**b**) Entropy of the system response.

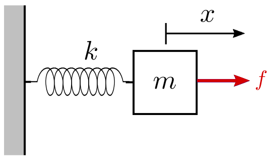


Figure 12: Forced mass-spring oscillator.

7.5 Example 5: forced vibration

Consider a simple mass-spring oscillator moving on a horizontal surface without friction subject to a force f, as shown in Fig. 12.

Its equation of motion is

$$m\, \ddot{x}(t) + k\, x(t) = f(t)\,. \tag{20}$$

The driving force f is chosen to be of the form

$$f(t) = t\,, \tag{21}$$

and the initial conditions are $x(0) = 1$ m and $\dot{x}(0) = 0$ m/s. Hence, the solution of Eq. 20 is

$$x(t) = \cos\left(\sqrt{k}t\right) + \frac{t}{k}\,. \tag{22}$$

Consider again that the spring stiffness is uncertain and modeled as a discrete random variable K with Bernoulli distribution given by Eq. 16. The position of the mass becomes a random process, $\mathcal{X}$. The mean, variance and entropy of $\mathcal{X}$ are

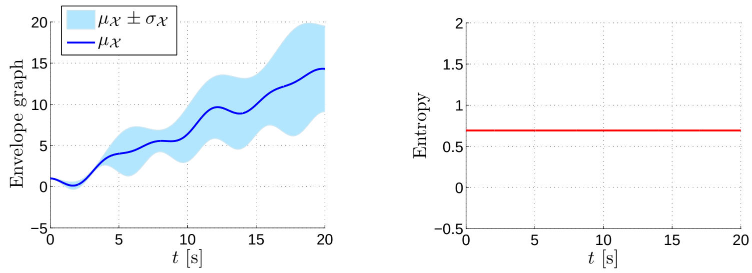
$$\mu_{\mathcal{X}}(t) = E[\mathcal{X}(t)] = \frac{1}{2}\left(\cos(t) + t + \cos\left(\frac{\pi}{2}t\right) + \frac{4t}{\pi^2}\right), \tag{23}$$

$$\begin{aligned}\sigma^2_{\mathcal{X}}(t) =& E[(\mathcal{X}(t) - \mu_{\mathcal{X}}(t))^2] \\ =& \frac{1}{2}\left\{\frac{\cos(t)}{2} + \frac{t}{2} - \frac{\cos\left(\frac{\pi t}{2}\right)}{2} - \frac{2t}{\pi^2}\right\}^2 + \frac{1}{2}\left\{\frac{\cos\left(\frac{\pi t}{2}\right)}{2} + \frac{2t}{\pi^2} - \frac{\cos(t)}{2} - \frac{t}{2}\right\}^2.\end{aligned} \tag{24}$$

$$\eta_{\mathcal{X}}(t) = -\frac{1}{2}\ln\left(\frac{1}{2}\right) + \frac{1}{2}\ln\left(\frac{1}{2}\right) \approx 0.69\,, \quad \forall t > 0\,. \tag{25}$$

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar

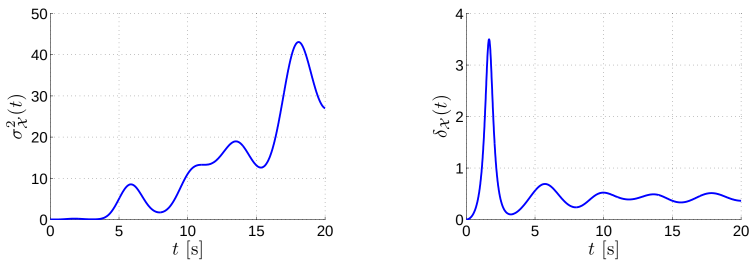
In Fig. 13(a) it is shown an envelope graph of the system response. The shaded region is bounded by mean ± standard deviation. The entropy, variance and coefficient of variation graphs are shown in Fig. 13(b), 14(a) and 14(b).



(a) (b)

Figure 13: (**a**) Envelope graph of the system response. (**b**) Entropy of the system response.

The result shows that the entropy remains constant in time, despite the variation of the moments. Therefore, the envelope graph suggests that the "uncertainty" is growing with time while the entropy suggests that it is constant. The variance and coefficient of variation as function of time, suggest other behaviors. While the variance oscillates over time, the coefficient of variation has a peak between zero and five seconds. This example shows that the three strategies found in the literature are not compatible. They provide different information and can mislead us and lead to errors.



(a) (b)

Figure 14: (a) Variance of the system response as function of time. (b) Coefficient of variation of the system response as function of time.

7.6 Example 6: stochastic stick-slip oscillations

The previous examples were chosen to be as simple as possible in order to highlight different problems that can arise when one uses the sets of statistics. Next we present a simple, but more realistic example. This example is discussed in details in the papers Lima and Sampaio (2017a,b).

Consider the system composed by a simple oscillator (mass-spring) moving on a rough surface, as shown in Fig. 15. The base has velocity v. The roughness induces a dry-frictional force between the mass and the base which is modeled as a Coulomb friction. Due to this friction model, the resulting motion of the mass can be characterized in two qualitatively different modes:

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar

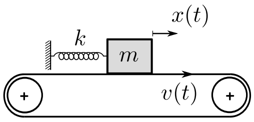


Figure 15: Stick-slip oscillator.

- the stick-mode (in which the mass and base have the same velocity during an open time interval) and;
- the slip-mode, in which mass and base have different velocities.

Consider that the dry-friction oscillator has an imposed stochastic bang-bang motion. Its velocity is modeled as a Poisson process, represented by $\mathcal{V}$.

Since it is assumed that the base motion is uncertain, the response of the stochastic stick-slip oscillator is a random process which presents a sequence alternating stick and slip-modes. We are interested in the stochastic characterization of these sequences. Defined a time interval for analysis, the variables of interest are the number of time intervals in which stick and slip occur, the instants at which they start and their duration. These variables are modeled as stochastic objects in order to allow the stochastic characterization the dynamics of the oscillator. Thus we have the

- number of time intervals in which stick occurs represented by the discrete random variable S_T;
- number of time intervals in which slip occurs represented by the discrete random variable S_L;
- instants at which sticks begin represented by a discrete random process $T_1, \cdots, T_{S_T}$, where the subscripts $1, \cdots, S_T$ indicate the order that they occur, i.e., the instant in which starts the first stick, the second, and so on up to the S_T-th stick;
- duration of sticks represented by a discrete random process $D_1, \cdots, D_{S_T}$, where again the subscripts $1, \cdots, S_T$ indicate the order that they occur;
- instants at which slips begin represented by a discrete random process $L_1, \cdots, L_{S_L}$, where $1, \cdots, S_L$ indicate the order that they occur;
- duration of sticks represented by a discrete random process $H_1, \cdots, H_{S_L}$, where $1, \cdots, S_L$ indicate the order that they occur.

Figure 16 shows a sketch of the sequence of sticks and slips in the system response. Observe that we count the first slip just after the first stick, i.e., we have $L_1 > T_1$. Besides this, if the chronometer stops during a slip, the number of sticks is equal or the number of slips, i.e. $S_T = S_L$. If the the chronometer stops during a stick, then $S_T = S_L + 1$.

To estimate statistics and histograms of the random variables that characterize the system response, the dynamical equations were integrated 18,000 times using independent realizations of the base movement generated with the Monte Carlo method Sampaio and Lima (2012); Souza de Cursi and Sampaio (2015). Details of the numeric simulations can be found in Lima and Sampaio (2017a,b).

Some estimated statistics, variance, coefficient of variation, and Shannon entropy of $T_1, \cdots, T_6$, $D_1, \cdots, D_6$, $L_1, \cdots, L_6$ and $H_1, \cdots, H_6$ were computed. These results are shown graphically

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar

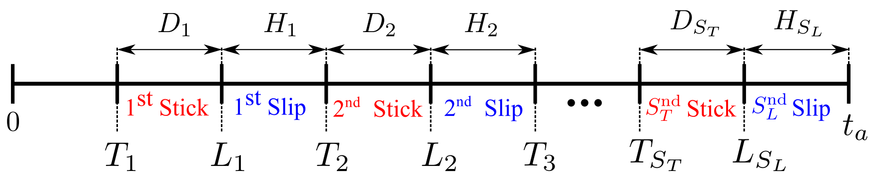
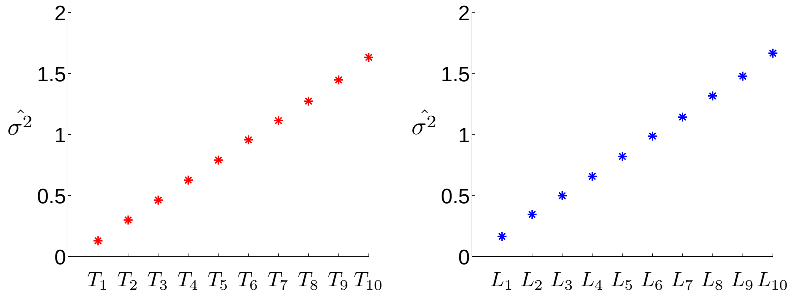


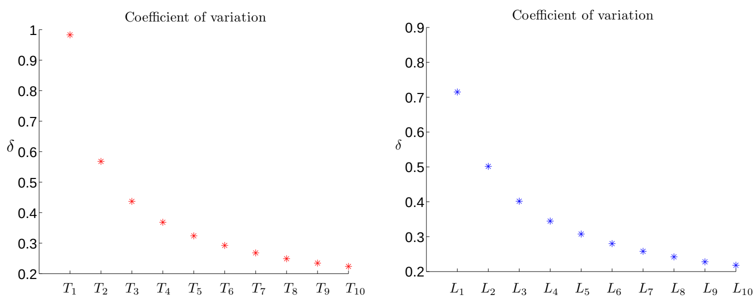
Figure 16: Sketch of sequence of sticks and slips in the system response for the case in which $S_T = S_L$.

in Figs. 17(a) to 19(b). It can be observed that the variance and Shannon entropy of the instants at which the sticks and slips start grows with the stick and slip number. However, the coefficient of variation decreases. This example shows that the strategies found in the literature provide



(a) (b)

Figure 17: Estimated variance of the instants at which the(a) sticks and (b) slips start.



(a) (b)

Figure 18: Estimated coefficient of variation of the instants at which the(a) sticks and (b) slips start.

different information in relation to the behavior of uncertainty. They are not compatible. To characterize uncertainty propagation, it is necessary to determine the variation of the CDF, as it is done in Lima and Sampaio (2017a,b). In these two papers, marginal and joint densities functions of the instants at which sticks and slips start and their duration are investigated.

8 CONCLUSIONS

In this paper we define what is, and what is not, uncertainty quantification and propagation. Uncertainty is described by the cumulative distribution function (CDF) and, uncertainty propagation is the variation of the cumulative distribution function (CDF) with a parameter, which can be discrete or continuous. It is not the variation of sets of statistics one with a parameter.

To replace the CDF for a small set of statistics seems to be a powerful tool to determine uncertainty propagation. However, it can lead to errors and misleadingness. With simple examples, we show that the strategies based on a set of statistics may not be compatible, moreover,

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar

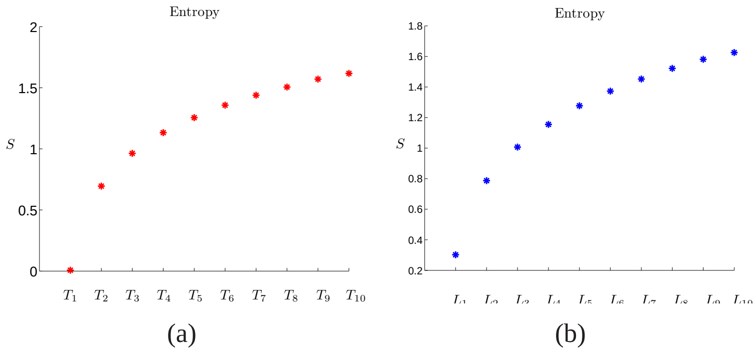


Figure 19: Estimated Shannon entropy of the instants at which the(a) sticks and (b) slips start.

they can be even contradictory. Regarding the examples one sees immediately the inadequacy of the use of reductions to compute uncertainty propagation. Furthermore, some of the strategies based on a set of statistics can not always be used. Some random variables do not have any moment, that is, it is not possible to compute neither envelop nor coefficient of variation for them.

In the literature, it is possible to find papers that use simultaneously, two or more of these different sets of statistics to measure of uncertainty propagation without realizing that they are not compatible. Examples are the works that try to characterize how the uncertainty of an inputs of a system affects the system response. For instance, the Maximum Entropy Principle is used to construct the probability model of the input with the argument that it maximizes the uncertainty. After, for different values of the coefficient of variation of the input statistics of the system response are computed. With these statistics, an envelope graph is construct. The objective is to establish a relation between the value of the coefficient of variation of the input with envelope graph of the output.

The expressions “uncertainty” and “uncertainty quantification” are fashionable and largely used in the literature. However, uncertainty is the CDF. The sets of statistics used so far, can not replace the CDF. Moreover, since the sets of statistics we described are not compatible, they should not be mixed in the same paper. Besides, if one uses one of these sets, clearly specified, we see no need to call it “uncertainty”. There are better description, variance measures absolute error, coefficient of variation relative error, and entropy is entropy, no new names are needed! As Jaynes said, the things should be called by their names, and no new names should be invented Jaynes (1978).

ACKNOWLEDGEMENTS

The authors acknowledge the support given by FAPERJ, CNPq and CAPES.

REFERENCES

Chen J., Eeden C., and Zidek J. Uncertainty and the conditional variance. *Statistics and Probability Letters*, 80:1764–1770, 2010.

Chung K. *A course in probability theory*. Academic Press, 1974.

Conrad K. Probability distributions and maximum entropy. *http://www.math.uconn.edu/ kconrad/blurbs/analysis/entropypost.pdf*, Retrieved July 24:1–27, 2016.

Einstein A. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, 322(8):549–560, 2018. doi:10.1002/andp.19053220806.

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar

Feller W. *An Introduction to Probability Theory and its Applications*, volume I and II. Wiley, 1957.

Ghanem R., Higdon D., and Owhadi H. *Handbook of Uncertainty Quantification*. Springer, 2017.

Grimmett G. and Welsh D. *Probability an introduction*. Oxford science publications, United States, New York, 1986.

Jaynes E. Information theory and statistical mechanics. *The Physical Review*, 106(4):620–630, 1957.

Jaynes E. Where do we stand on maximum entropy? In *The Maximum-Entropy Formalism*. Papers on Probability, Statistics and Statistical Physics, Massachusetts Institute of Technology (MIT), 1978.

Khodabin M. and Ahmadabadi A. Some properties of generalized gamma distribution. *Mathematical Sciences*, 4(1):9–28, 2010.

Khosravi A. and Nahavandi S. An optimized mean variance estimation method for uncertainty quantification of wind power forecasts. *Electrical Power and Energy Systems*, 61:446–454, 2014.

Le Maître O. and Knio O. *Spectral Methods for Uncertainty Quantification*. Springer, United States, New York, 2010.

Lima R. and Sampaio R. Construction of a statistical model for the dynamics of a base-driven stick-slip oscillator. *Mechanical Systems and Signal Processing*, 91:157–166, 2017a.

Lima R. and Sampaio R. Parametric analysis of the statistical model of the stick-slip process. *Journal of Sound and Vibration*, 397:141–151, 2017b.

Motra H., Hildebrand J., and Wuttke F. The Monte Carlo method for evaluating measurement uncertainty: Application for determining the properties of materials. *Probabilistic Engineering Mechanics*, 45:220–228, 2016.

Nordström J. and Wahlsten M. Variance reduction through robust design of boundary conditions for stochastic hyperbolic systems of equations. *Journal of Computational Physics*, 282:1–22, 2015.

Sampaio R. and Lima R. *Modelagem Estocástica e Geração de Amostras de Variáveis e Vetores Aleatórios*, volume 70 of *Notas de Matemática Aplicada*. SBMAC, http://www.sbmac.org.br/arquivos/notas/livro_70.pdf, 2012.

Shannon C. A mathematical theory of communication. *Bell System Tech.*, 27:379–423 and 623–659, 1948.

Smith R. *Uncertainty Quantification: Theory, Implementation, and Applications*. SIAM, 2014.

Soize C. *Stochastic Models of Uncertainties in Computational Mechanics*. American Society of Civil Engineers, 2012.

Soize C. *Uncertainty Quantification: An Accelerated Course with Advanced Applications in Computational Engineering*. Springer, 2017.

Souza de Cursi E. and Sampaio R. *Uncertainty Quantification and Stochastic Modeling with Matlab*. Elsevier, ISTE Press, 2015.

Sullivan T. *Introduction to uncertainty quantiffication*. Springer, United States, New York, 2015.

Zidek J. and Eeden C. *Uncertainty, Entropy, Variance and the Effect of Partial Information*, volume 42. Lecture Notes-Monograph Series, Institute of Mathematical Statistics, 2003.

Copyright © 2018 Asociación Argentina de Mecánica Computacional http://www.amcaonline.org.ar